VITAL AND HEALTH STATISTICS

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES • Public Health Service • National Center for Health Statistics • Series 1, No. 16

HE 2

nchs

# Data Systems of the National Center for Health Statistics

Programs and Collection Procedures
Series 1, No. 16

Library of Congress Cataloging in Publication Data

National Center for Health Statistics (U.S.)
Data systems of the National Center for Health Statistics.

(Vital and health statistics. Series 1, Programs and collection procedures ; no. 16) (DHHS publication ; no. (PHS) 82-1318)
Prepared by: Nancy D. Pearce.
Supt. of Docs. no.: HE 20.6209:1-16
1. National Center for Health Statistics (U.S.) 2. Public health—United States—Statistical services. 3. Medical care—United States—Statistical services. 4. United States—Statistics, Medical. I. Pearce, Nancy D. II. Title. III. Series. IV. Series: DHHS publication ; no. (PHS) 82-1318. [DNLM: 1. Data collection. 2. Health surveys. United States. W2 A N148vs no. 16]
RA409.U44 no. 16 [RA407.3] 312'.0723s 81-607814
ISBN 0-8406-0232-4 AACR2

VITAL AND HEALTH STATISTICS

# Data Systems of the National Center for Health Statistics

Each of the data collection programs of the National Center for Health Statistics is described. For each program the background, purpose, scope, sample design, data collection procedures, content of data collection forms, method of data release, and uses of data are presented.

Programs and Collection Procedures
Series 1, No. 16

DHHS Publication No. (PHS) 82-1318

U.S. Department of Health and Human Services
Public Health Service
Office of Health Research, Statistics, and Technology
National Center for Health Statistics
Hyattsville, Md.
December 1981

**National Center for Health Statistics**

DOROTHY P. RICE, *Director*

ROBERT A. ISRAEL, *Deputy Director*

JACOB J. FELDMAN, Ph.D., *Associate Director for Analysis and Epidemiology*

GAIL F. FISHER, Ph.D., *Associate Director for the Cooperative Health Statistics System*

GARRIE J. LOSEE, *Associate Director for Data Processing and Services*

ALVAN O. ZARATE, Ph.D., *Assistant Director for International Statistics*

E. EARL BRYANT, *Associate Director for Interview and Examination Statistics*

ROBERT C. HUBER, *Associate Director for Management*

MONROE G. SIRKEN, Ph. D., *Associcte Director for Research and Methodology*

PETER L. HURLEY, *Associate Director for Vital and Health Care Statistics*

ALICE HAYWOOD, *Information Officer*

READEX FICHE

best copy available

# Foreword

This brief, semitechnical description of the various data systems of the National Center for Health Statistics has been created to serve a variety of needs. The greatest need is for a single publication to bring together different descriptive materials under one cover. For the user who needs more detailed, technical knowledge of these systems, this is available in separate publications for the individual programs. It is expected that this summary publication will serve the needs of many visitors to the Center, as well as the needs of those at the Federal, State, and local levels who wish to understand the structure and scope of Center programs.

One feature that should be carefully noted is the flexible nature of many of these surveys. While subject to the need for standard measurement over time, the surveys are also able to change with the demands for new data. Systematic reviews lead to careful revisions and additions as needs are identified. When necessary, new survey activities are designed and fielded. Thus, this report is a cross-sectional description of the current data collection program of the Center.

Dorothy P. Rice
Director

blank page

# Contents


Foreword . . . . . iii

Introduction . . . . . 1
- Background . . . . . 1
- Confidentiality of data . . . . . 1
- Data collection mechanisms . . . . . 2
- Morbidity and mortality data coding . . . . . 2
- Data release mechanisms . . . . . 2
- Descriptive, methodological, and other related reports . . . . . 3

Vital statistics . . . . . 4
- Basic vital statistics . . . . . 4
- Vital statistics followback surveys . . . . . 6
- National Survey of Family Growth . . . . . 8

General population surveys . . . . . 11
- National Health Interview Survey . . . . . 11
- National Medical Care Utilization and Expenditure Survey . . . . . 14
- National Health and Nutrition Examination Survey . . . . . 16

Health resources utilization statistics . . . . . 20
- National Hospital Discharge Survey . . . . . 20
- National Ambulatory Medical Care Survey . . . . . 21
- National Nursing Home Survey . . . . . 23

Health resources statistics . . . . . 26
- National Master Facility Inventory . . . . . 26
- National health professions inventories and surveys . . . . . 27

**Appendixes**

Contents . . . . . 30
I. Definition of certain terms used in this report . . . . . 31
II. Listing of reports related to specific programs . . . . . 34

**List of text tables**

A. Sample size and total response rate (percent) for each of the vital statistics followback surveys: United States, 1961-80 . . 7
B. Response rates for National Nursing Home Surveys, by type of questionnaire used (percent): United States, 1973-74 and 1977 . . . . . 25

## Symbols

| | |
|---|---|
| - - - | Data not available |
| . . . | Category not available |
| - | Quantity zero |
| 0.0 | Quantity more than zero but less than 0.05 |
| Z | Quantity more than zero but less than 500 |
| * | Figure does not meet standards of reliability or precision (more than 30-percent relative standard error) |
| # | Figure suppressed to comply with confidentiality requirements |

# Data Systems of the National Center for Health Statistics

by Nancy D. Pearce, Office of the Center Director

## Introduction

### Background

The National Center for Health Statistics is one of the major Federal statistical organizations. It operates a diverse survey and inventory program with legislative authorization to collect statistics on the following:

- The extent and nature of illness and disability of the population of the United States, including life expectancy, maternal morbidity, and mortality.
- The impact of illness and disability of the population on the economy of the United States and on other aspects of the well-being of its population.
- Environmental, social, and other health hazards.
- Determinants of health.
- Health resources, including health professionals by specialty and type of practice; and the supply of services by hospitals, extended care facilities, home health agencies, and other health institutions.
- Utilization of health care, including utilization of ambulatory health services, services of hospitals, extended care facilities, home health agencies, and other institutions.
- Health care costs and financing.
- Family formation, growth, and dissolution.

This report provides brief descriptions of each of the data systems operated by the National Center for Health Statistics. Each project description contains a statement of the purpose of the data system, the contents of the data collection forms, an overview of the data collection process, and an indication of the availability of the data to the public. Although this report is nontechnical in nature, it does include some technical terms, a limited glossary of which is included as appendix I. For the purposes of many of the Center surveys, however, many terms have highly specialized definitions. No effort has been made here to provide a comprehensive series of such definitions. Each analytic report in series 10 through 23 (see "Data release mechanisms") includes in its own appendix the definitions of certain specialized terms used.

### Confidentiality of data

Data collection activities of the Center are subject to the specific provisions of both the Privacy Act of 1974 (Public Law 93-579) and the Health Services Research, Health Statistics, and Health Care Technology Act of 1978 (Public Law 95-623). The Privacy Act covers all federally sponsored and operated data collection that involves creation of a system of records containing unique personal identifiers, while the latter constitutes legislative authorization for a portion of the Public Health Service.

In keeping with the requirements of the Privacy Act of 1974 and Center policy, each individual, household, or establishment asked to provide data for the Center is informed in writing of the following:

- The authorization for soliciting the information.
- That disclosure of such information is voluntary.
- The principal purpose or purposes (usually statistical or research) for which the information is intended to be used.
- The routine uses (as published in the *Federal Register*) which may be made of the information.
- The effects on the recipient, if any, of not providing all or any part of the requested information.

The Center is also bound by the provision of its authorizing legislation that the information obtained in the course of its statistical activities may not be

used for any purpose other than that for which it was supplied. Such information may not be published or released in other form if the particular establishment or person supplying the information or described in it is identifiable, unless such establishment or person has consented to its publication or release in other form. Thus, the Center assurance to respondents that data will be kept confidential has a legal basis.

Almost all data collected by the Center are requested under a pledge to the respondent (either an individual or an establishment) that the information will be used only for statistical purposes and will not be released in such a way that individual persons, households, or facilities will be identifiable. The Center never releases information that would identify an individual or a household. Some information is released that identifies individual facilities, however. When it is planned to release such information, facility administrators are informed at the time they provide the data which, if any, items will not be accorded confidential treatment.

The *NCHS Policy Statement on Release of Data for Individual Elementary Units and Special Tabulations* presents and explains in detail the confidentiality policy.

### Data collection mechanisms

The Center does not have a large data collection staff of its own. It collects most of its data through interagency agreements with the U.S. Bureau of the Census or through contracts with non-Federal organizations.

It has, for a number of years, been Center policy to increase the capacity at the State level to collect and utilize data. As in the case of vital statistics data, the Center also tries to identify instances when it is possible to unduplicate the collection of data by having it collected at the most appropriate point (Federal, State, or local) and then making it available to others who have need for it. While vital statistics is currently the only area in which the Center relies on this method of data collection and sharing for the production of national data, it is hoped that in the years to come other types of data may be collected through a national data-gathering system operating on the State level, the Cooperative Health Statistics System.

### Morbidity and mortality data coding

The Center collects a wide range of information about illness and the health problems of the population of the United States. In order to enhance the comparability of international statistics and of morbidity and mortality estimates from such diverse sources as household respondents and health care provider records, these data are coded to appropriate adaptations of the International Classification of Diseases (ICD). ICD is a system for classification of diseases and injuries for statistical purposes and is revised under the auspices of the World Health Organization every 10 years.

All Center mortality and morbidity information is classified according to the most recent revision of ICD. Data for the approximately 10-year period prior to January 1, 1979, are coded to the *Eighth Revision International Classification of Diseases, Adapted for Use in the United States* (ICDA-8). Mortality data for the period beginning in January 1979 are classified according to the ninth revision of the International Classification of Diseases (ICD-9), and morbidity data are classified using the clinical modification of ICD-9 (ICD-9-CM).

### Data release mechanisms

The Center releases its data in several ways. The *Vital and Health Statistics* series contain detailed reports on the design of the various data collection systems, findings of methodological studies, and detailed cross-tabulations of data. *Advance Data* reports are timely but brief (4-12 pages each) summaries of newly available data on topics of special interest. These reports have replaced supplements to the *Monthly Vital Statistics Report* for all survey programs except the natality and mortality follow-back surveys. All these reports are distributed without charge, as issued, to libraries, medical colleges, schools of public health, and other selected institutions that have requested to receive all issues in the series. Other organizations and individuals receive publication notices and information to purchase copies from the Government Printing Office. The annual *Catalog of Publications of the National Center for Health Statistics* contains a cumulative list of these reports, and the quarterly *Publication Note* is issued to provide the latest information.

The Center also releases an annual booklet that describes public use data tapes available for purchase. This program, begun in 1969, permits researchers to perform their own analyses of data. Tape availability and contents, along with information on ordering the tapes, are detailed in each year's *Catalog of Public Use Data Tapes From the National Center for Health Statistics.* The majority of these data tapes are now sold by the National Technical Information Service (NTIS) and are not available from the Center. Eventually, all Center public use data tapes will be available from NTIS. The order form in the back of the catalog gives the address and instructions for ordering from NTIS or, in a few instances, from the National Center for Health Statistics.

Many requests for unpublished data are filled regularly using tabulations that have been previously compiled. In addition, many special tabulations are

prepared each year to meet data requests that cannot be fulfilled in any other way. When special tabulations are necessary, the requester is given a cost and time estimate. Tabulations and public use data tapes are carefully reviewed to ensure that confidentiality is maintained.

### Descriptive, methodological, and other related reports

In addition to the reports from each program that are issued in the analytic series designated in the "release of data" section in each program description, there are reports in other *Vital and Health Statistics* series that are directly related to the programs. These include program descriptions, methodological reports, and reports that are in-depth analysis or contain policy recommendations. A listing of these reports, by program, is presented in appendix II.

Requests for publications and information or inquiries concerning data tapes, special tabulations, and other assistance should be directed to the Scientific and Technical Information Branch.

# Vital statistics

## Basic vital statistics

*Background.*—Basic vital statistics come from records of live births, deaths, fetal deaths, induced terminations of pregnancy, marriages, and divorces or dissolutions of marriages. Registration of these events is a local and State function, but uniform registration practices and use of the records for national statistics have been established over the years through cooperative agreements between the States and the National Center for Health Statistics and its predecessor agencies.

*Purpose and scope.*—The purpose of the basic vital statistics program is to formulate and maintain a cooperative and coordinated vital records and vital statistics system, promoting high standards of performance. The program is nationwide in scope, covering the entire population of the United States.

*Data collection procedures.*—Both provisional and final vital statistics are derived from the registration system. The provisional data are obtained from counts of vital records registered without reference to the date the event occurred; the final data are obtained from the record and its contents, processed by date of occurrence of the event.

The civil laws of every State provide for a continuous and permanent birth, death, and fetal death registration system. In general, the local registrar of a town, city, county, or other geographic place collects the records of births and deaths occurring in the area; inspects, queries, and corrects if necessary; maintains a local copy, register, or index; and transmits them to the State health department. There the vital statistics office inspects the records for promptness of filing and for completeness and consistency of information; queries if necessary; numbers, indexes, and processes the statistical information for State and local use; and binds the records for permanent reference and safekeeping. Microfilm copies of the individual records or machine-readable data are transmitted to the National Center for Health Statistics for use in compiling the final annual national vital statistics volume.

The system for collecting national data on marriages and divorces is not as well developed as the system for births and deaths is. All States have marriage and divorce laws; but, as of January 1, 1981, three States did not have a central file of marriage records, and four States did not have a central file of divorce records. Forty-two States and the District of Columbia are in the designated marriage-registration area, and 30 States are in the divorce-registration area. These registration-area States provide the Center microfilm copies of their marriage and divorce certificates or machine-readable data from which the Center derives data on characteristics of marriages and divorces and the persons involved. States not in the registration areas provide only counts of marriages by month and counts of marriages and divorces by county. For States without central registration, final counts are collected from individual counties either by the State vital statistics office or by the Center. In these instances some counties report only marriage licenses issued and divorce petitions filed rather than marriages performed and divorces granted.

Most States submit microfilm copies of vital certificates or machine-readable data tapes prepared from the certificates to the National Center for Health Statistics within 90 days following the end of each data month. Special arrangements for less frequent shipments are made with smaller registration areas. Relevant information derived from answers to State queries about the original certificates and received by the State after shipment of the certificates to the Center are filmed and submitted with the first shipment following their receipt by the State.

Provisional vital statistics are collected and published monthly and summarized annually. They are derived from monthly reports from the States to the Center giving the number of certificates accepted by

each State for filing between two dates a month apart, without regard to actual date of occurrence. These reports to the Center are mailed on or before the 25th of the month following the data month. They are the source of the provisional vital statistics published in the *Monthly Vital Statistics Report* and its Annual Summary. Provisional data also include a 10-percent sample of death certificates, the Current Mortality Sample, which provides provisional cause-of-death data on a monthly basis. The Current Mortality Sample is selected by the Center from the regular data file of deaths for those States submitting their entire month's file by the end of the following month. Otherwise the State is asked to provide a sample of records on a current basis. The sample is selected by including each record with a given last digit in the certificate number.

*Collection forms.*—To promote uniformity in the statistical information collected from States and local areas for national statistical purposes, the National Center for Health Statistics recommends standard certificates or reports for birth, death, fetal death, induced termination of pregnancy, marriage, and divorce. The standard certificates and reports are developed cooperatively with the States and local areas and the Center and its predecessor agencies, taking into account the needs and problems expressed by the major providers and users of the data. They are reviewed about every 10 years to ensure that they meet to the fullest extent feasible current needs as legal records and as sources of vital and health statistics. Although the use of standard certificates and reports by States is voluntary, and their form and content may vary according to the laws and practices of each State, the certificates and reports in most States closely follow the standard.

The Series 4 report "The 1968 Revision of the Standard Certificates" outlines the history and basic principles of the standard certificates and the major objectives of the 1968 revisions; the report also contains copies of the standard certificates and describes the principal additions, modifications, and deletions of items. A similar report in preparation describes the 1978 revision of the standard certificates.

*Data processing methods.*—Machine-readable data are accepted from States participating in the Cooperative Health Statistics System. Data from all other States are coded and entered on magnetic tape.

All death and fetal death records are processed. For States not participating in the Cooperative Health Statistics System, a 50-percent sample of live births is processed by selecting records with image numbers terminating in an even number. For States in this cooperative system the entire birth file is used. In general, all States in the marriage and divorce registration areas that do not submit these data under the system send microfilm copies of all their marriage or divorce certificates to the Center, where a sample of the records is selected and processed.

*Release of data.*—Vital statistics data reach the public through the *Monthly Vital Statistics Report* and its supplements, *Vital Statistics of the United States* (bound volumes issued annually), public use data tapes (issued annually), and *Vital and Health Statistics* series reports and special reports. Data on mortality are published in Series 20; and data on natality, marriage, and divorce are published in Series 21. Unpublished tabulations may also be available upon request.

The *Monthly Vital Statistics Reports* containing provisional monthly counts are published within about 2 months following the end of the data month. Provisional cause-of-death information from the Current Mortality Sample are published about 3 months after the data month, and the *Monthly Vital Statistics Report* advance supplements releasing final annual statistics are published in about 18 months after the data year. Public use data tapes are released to the National Technical Information Service within 1 month after publication of the advance supplements; series and special reports are issued as resources permit; and the annual *Vital Statistics of the United States* are soon expected to be distributed with about a 36-month lag.

For vital records, names and addresses are not coded and never appear on data tapes or tabulations. Certificate numbers are never released without written permission from the States whose records are involved, and data tapes including data from localities having less than 250,000 population are classified to reflect population size without revealing specific geographic areas.

*Uses of data.*—When vital statistics data have been collected, processed, and made available for public use, they serve a variety of purposes. They are used in making population estimates and projections, in assessing the health of the U.S. population, in pinpointing health problems, and in measuring progress made by national health programs. They are also used in epidemiological studies; in marketing research; in demographic, sociological, and economic studies; in studies of the family; and in other numerous and varied aspects of research aimed at understanding our society and its problems and progress.

Among the principal users of the data are other parts of the Public Health Service, such as the Centers for Disease Control, the Food and Drug Administration, and the National Institutes of Health; other Federal agencies, such as the Bureau of the Census, Federal Reserve Board, Bureau of Labor Statistics, and Departments of Agriculture, Defense, and Transportation; members of the United States Congress; State and city governments; public and private research institutions; life insurance companies; the

faculty and students of universities; newspaper reporters and feature writers; and physicians and workers in health information and education groups.

### Vital statistics followback surveys

*Background.*—National natality and mortality surveys are periodic data collections based on samples of registered deaths and births occurring during a calendar year. Mortality surveys were conducted annually from 1961 through 1968, and natality surveys from 1963 through 1969 and again in 1972. A National Infant Mortality Survey was also conducted from 1964 to 1966. For 1980, a National Natality Survey was conducted that was expanded to include fetal mortality.

*Purpose and scope.*—The national followback surveys extend for statistical purposes the range of items normally included on the vital records. They provide national estimates of births and deaths by characteristics not available from the vital registration system. They also serve as a basis for evaluating the quality of information reported on the vital records.

*Sample design.*—The birth or death record serves as the sampling unit, and samples of these units are selected from a frame of records representing births or deaths registered during a given period, usually a calendar year.

The sampling frame for the National Mortality Survey is the Current Mortality Sample, the 10-percent systematic sample of death certificates received each month by the National Center for Health Statistics from the registration areas in the United States. The sample for the National Mortality Survey is a sample subselected monthly from the Current Mortality Sample. The sampling frame for the National Natality Survey is the file of birth certificates from each of the registration areas of the United States. Each registration area assigns a file number to each birth certificate, and these file numbers run consecutively from the first to the last birth occurring in that area during that year. The samples for the 1980 survey are based on a probability design that makes use of these certificate numbers. For the sample of live births each 500th birth certificate record was systematically selected; for the sample of fetal deaths each 4th fetal death certificate was systematically selected. The sample size and the total response rate for each of the surveys are shown in table A.

*Data collection procedures.*—Data for all the followback surveys are collected primarily by mail. In the natality surveys, from addresses given on the birth certificates, questionnaires are sent to the mother, the physician who delivered the baby, and the medical facility where the baby was born. In addition, the 1980 surveys obtained data from the medical sources who the woman named as having given her radiation treatment or examination in the year prior to the delivery.

In the mortality surveys, a questionnaire is sent to the person who provided the funeral director with the decedent's personal information for recording on the certificate. This questionnaire requests socioeconomic information about the decedent as well as the names and addresses of hospitals and institutions that might have provided care to the decedent at any time during his or her last year of life. If the death occurred in a hospital or institution, a hospital questionnaire is sent directly to the hospital or institution asking for information about the care provided and for the names and addresses of other medical facilities providing care.

*Content of questionnaires.*—The questionnaires for past National Mortality Surveys have contained questions concerning the patient's last year of life. The 1961 survey included questions on hospital utilization, diagnoses, operations performed, institutions in which hospitalized, number of hospital stays, place of death, income, and whether working or retired during most of the last year of life. The 1962-63 questionnaire included a few questions on hospital and institutional care in the last year of life, household composition, education, and income, with detailed questions on place of residence of the decedent. The 1964 and 1965 questionnaires included questions on health insurance, place of hospital and institutional care, the charges of hospital care and who paid them, operations performed, how much of the surgeon's bill was paid by insurance, household composition, assets, and income. The 1966-68 survey also included questions on the smoking habits of the deceased and questions regarding the household and income.

The questionnaire for the 1964-66 National Infant Mortality Survey included questions on hospitalization of the infant who died, information about other children of the mother, household composition, income, employment of mother, education of mother and father, and health insurance.

The National Natality Surveys collect information from mothers who had live births during a given year. The 1963 survey gathered information on the medical and dental care and radiological treatment of the mother, employment and education of mother and father, and family income. The 1964-66 survey gathered information on pregnancy history, expectations of having more children, household composition, income, whether this was a first or later marriage, date of first marriage and date of this marriage, whether mother was employed and when during her pregnancy she stopped working, education, and health insurance. The 1967-69 survey dealt with the medical care and smoking habits of the mother, breast feeding, information on earlier pregnancies, expectations of having more children,

Table A. Sample size and total response rate (percent) for each of the vital statistics followback surveys: United States, 1961-80

| Survey | Sample size | Response rate |
|---|---|---|
| National Mortality Survey, 1961: | | |
| Hospitals and institutions | 1 out of 330 deaths | 98 |
| Informants providing names of hospitals and institutions | 1 out of 330 deaths | 93 |
| National Mortality Survey, 1964-65 | 1 out of 330 deaths | 91 |
| National Infant Mortality Survey, 1964-65 | 1 out of 110 infant deaths | 88 |
| National Natality Survey, 1963: | | |
| Mothers | 1 out of 1,000 births | 86 |
| Physicians | 1 out of 1,000 births | 93 |
| Medical facilities | 1 out of 1,000 births | 98 |
| Dentists | 1 out of 1,000 births | 97 |
| National Natality Survey, 1964-66 | 1 out of 1,000 births | 89 |
| National Natality Survey, 1967 | 1 out of 1,000 births of white infants and 1 out of 500 births of all other infants | 91 |
| National Natality Survey, 1968-69 | 1 out of 1,000 births of white infants and 1 out of 500 births of all other infants | 87 |
| National Mortality Survey, 1966-68 | 1 out of 260 of all deaths of persons 35-84 years of age | 92 |
| National Natality Survey, 1972: | | |
| Mothers | 1 out of 500 births | 72 |
| Physicians | 1 out of 500 births | 72 |
| Hospitals | 1 out of 500 births | 85 |
| National Natality Survey, 1980: | 1 out of 350 births, including an oversampling of low birth weight infants and 2 out of 5 fetal deaths of 28 weeks gestation or more | Provisional rates are all 75-80 |
| Mothers | | |
| Physicians | | |
| Hospitals and other medical sources providing X-ray and ultrasound or nuclear medicine | | |

whether this was her first marriage, dates of first and present marriages, education and religious preference of husband and wife, information on household composition, work history, and income. The 1972 survey covered sources of medical care, health status of mother and infant, health insurance, outcomes of all pregnancies, whether pregnancy was wanted and expectations of having more children, whether the mother had an operation to prevent future pregnancies, whether this was the first marriage, dates of first and present marriages, education of mother and father, information about the household, and family income.

The 1980 national natality survey was broadened to include fetal mortality. It included many of the same or similar questions as previous surveys in order to allow trend studies in the areas of smoking habits, marriage and pregnancy history, education, income, health status of mother and infant, sterilization, radiological treatment, employment, childbearing expectations, and breast feeding. Many new areas of study were added, such as alcohol consumption, electronic fetal monitoring, amniocentesis, additional maternal and infant health indicators, occupation of mother and father, and ethnicity. Furthermore, there was an oversampling of low birth weight infants in the natality survey, which will enable more in-depth study of this high risk group.

Copies of the data collection forms for the various surveys are included as an appendix in each analytic report based on data from a particular survey.

*Data processing methods.*—After all methods of obtaining complete questionnaires have been exhausted, the data are edited and coded. Basic range edits are edits made to eliminate some coding errors. Computer processing includes consistency checks, estimation or assignment of weights, and imputation of missing data.

*Release of data.*—As with the basic vital statistics provided through the registration system, the followback survey data are released through the *Vital and Health Statistics* series reports, supplements to the *Monthly Vital Statistics Report*, and public use data tapes, as the data become available. Series reports issued prior to 1975 were published in Series 22,

"Data from the National Natality and Mortality Surveys"; since 1975 reports from these sample surveys based on vital records have been included in Series 20, "Data on Mortality," and Series 21, "Data on Natality, Marriage, and Divorce," as appropriate. The privacy and confidentiality of study participants are carefully maintained.

*Uses of data.*—Data from the followback surveys are in demand by all major users of vital statistics data. They help meet the expanding needs for natality and mortality data in public health, medical research, epidemiology, demography, and other related fields. Data provided through the natality surveys are used by agencies and individuals responsible for or studying maternal and child health care programs, and by researchers and others studying social and demographic issues.

### National Survey of Family Growth

*Background.*—The National Survey of Family Growth (NSFG) is a multipurpose statistical survey that provides a wide range of information serving needs of persons and organizations concerned with the dynamics of population change, family planning, and maternal and child health. Developmental funds and necessary positions were provided for establishing NSFG in the National Center for Health Statistics in fiscal year 1971. Fieldwork for the first cycle of the survey was begun in July 1973 and completed in January 1974. The second cycle began in January 1976 and was completed in September 1976. A third cycle is expected to begin in January 1982.

*Purpose and scope.*—NSFG is designed to produce data on factors influencing trends and differentials in fertility, family planning practices, sources from which family planning advice and services are obtained, the effectiveness and acceptability of the various methods of family planning, and those aspects of maternal and child health that are most directly related to fertility and family planning. The survey is based on personal interviews with women of childbearing ages selected from a nationwide area probability sample of households and group quarters, exclusive of the institutional population and women living on military bases.

*Sample design.*—Cycles I and II of the National Survey of Family Growth were based on a cross-sectional sample of women in the conterminous United States, 15-44 years of age, who were married, who had been married, or who had never been married but who had offspring of their own living with them in the household. Excluded from the sample were women living in group quarters (that is, five or more unrelated persons 18 years or older, unrelated to the head of the household, who live and eat together, or six or more unrelated adults who live and eat together—as, for example, in college dormitories, barracks, or long-term care institutions) and never-married women without children.

NSFG is designed as a multistage area probability sample. In the first cycle of this survey, the first-stage primary sampling units included 101 standard metropolitan statistical areas, counties, parts of counties, and independent cities. Secondary sampling units consisted of enumeration districts or block groups within selected primary sampling units. Where feasible, secondary units were subdivided into third-stage listing units with a probability of selection proportional to estimated housing. In the fourth stage, housing units within the third-stage listing units were chosen by systematic sampling. A fifth level of sampling among eligible women within a household was required to obtain one and only one interview in each household with at least one eligible respondent. To ensure sufficient precision for presenting intragroup comparisons on various fertility variables, the black population was oversampled.

The overall sample design called for 10,000 completed interviews, including approximately 4,000 black women and 6,000 women of white and other racial groups. The response rate was an overall 81 percent. The final number of interviews was 9,797, consisting of 3,856 black respondents and 5,941 respondents of white and other races.

For Cycle II of NSFG a multistage area probability design based on 79 primary sampling units was developed. The remaining stages were broadly similar to those of the first cycle, except that a stratum for new housing (housing built since 1970), sampled from building permit listings, was added. Overall expected sample size remained at 10,000, with comparable racial composition. After a certain level of effort was completed in the fieldwork, remaining nonrespondents were subsampled at a rate of 1 in 2. The response rate for Cycle II was an overall 83 percent. The actual number of completed extended interviews was 8,611, composed of 3,009 black respondents and 5,602 respondents of white and other races.

For Cycle III the sample design is broadly the same as for Cycle II, but with adjustments to accommodate the more inclusive universe of women of any marital status, 15-44 years of age. The sample of 7,300 women will include women in group quarters such as college dormitories and boarding houses with more than five unrelated persons. As in the previous cycles of the survey, military bases and the institutional population will not be sampled.

*Data collection procedures.*—Data collection for NSFG is conducted by private contractors according to Center specifications. Personal, in-depth interviews are conducted with women identified as eligible for extended interviews by means of a household screening interview.

The questionnaires used for data collection in-

clude a household screener designed to obtain household composition data and to identify eligible extended-interview respondents in the sample households. For Cycles I and II, two different versions of the questionnaire were used for the extended interview—a "currently married" questionnaire for women who were married at the time of the interview and a "post-married" questionnaire for women who were widowed, divorced, or separated or who had never been married but had their own children living in the household. The differences between the two versions were a result primarily of rewording the questions and of deleting questions related to husbands in order to make the "post-married" questionnaire appropriate for respondents not married at the time of the interview. Throughout the questionnaires, interviewer instructions were used to skip respondents over questions, or entire questionnaire sections, that were not applicable to their individual situations. Cycle III of the survey involves a screener questionnaire similar to that used in the previous cycles and two separate questionnaires for the extended interviews, one for women under 25 years of age and one for women 25 years of age and over. The construction and format of the questionnaires are modeled on those of the previous cycles.

Standardized interviewer training programs, which provide in-depth training on the questionnaires and on NSFG concepts and procedures, are conducted in several different locations in the country. In addition to successfully completing training, interviewers are required to conduct, for review and approval, several practice interviews prior to beginning their field assignments. Only women are employed as interviewers and observers in NSFG. Field supervisors receive the same in-depth training as the interviewers, in addition to intensive training in field procedures.

Several quality control procedures were developed in Cycle I and refined in Cycle II to ensure the quality of the collected and processed data: a systematic field edit of selected interview items and a validation of a sample of each interviewer's work (a sample recheck) were performed throughout the fieldwork. Data preparation was validated through a 5-percent sample recode of all questionnaires, in addition to systematic verification of each coder's or keyer's work. A comprehensive legal-code and consistency cleaning program was developed and used for the data tapes. Similar quality control procedures are specified for Cycle III.

*Content of questionnaires.*—In Cycles I and II, the NSFG questionnaires covered marital history, a detailed pregnancy history, fecundity and expected or intended future births, pregnancy planning practices and utilization of specific contraceptive methods, the source and financing of family planning services, and a broad range of socioeconomic and demographic characteristics. Cycle III will collect information on sex education and the sexually active population as well as on the topics covered in Cycles I and II.

In the first cycle of the survey, the average length of an interview was 73 minutes; in the second cycle the average length of an interview was reduced to about 56 minutes; in the third cycle interviews will again be kept within an hour, on the average.

A copy of the Household Screener questionnaire for Cycle I is contained in appendix III of Series 2, No. 76; the screener for Cycle II is in appendix II of the comparable report for Cycle II, Series 2, No. 87. Relevant portions of the extended interview questionnaires are reproduced in appendixes of Series 23 reports. The "currently married" questionnaire for Cycles I and II is reproduced in appendix III of Series 4, No. 18.

*Release of data.*—The publication program for NSFG includes advance reports issued in *Advance Data,* methodological reports and in-depth substantive analyses in Series 2 and 23 of *Vital and Health Statistics,* and public use data tapes. Publications from the first cycle were issued beginning in 1976; the public use data tape for this cycle was released in 1977. Advance reports from the second cycle were issued beginning in 1978, and later the public use data tape became available in two parts: the first part, containing characteristics of respondents, including summary fertility characteristics, was released in April 1979; and the second part, containing characteristics relating to each birth, became available in 1980. A public use data tape for Cycle III is expected to be available by the spring of 1984.

*Uses of data.*—The rapidly growing need for collection of data on trends and differentials in the birth rate, family planning, and related aspects of maternal and child health has been recognized by the President and Congress. Attention has been directed toward studying population trends, assessing the implications of future population growth, making appropriate policy recommendations, and helping to educate the Nation regarding the consequences of population growth. More information is needed than is currently available from the regular data collection programs and occasional surveys of the U.S. Bureau of the Census and the National Center for Health Statistics. NSFG has been developed to provide, on a periodic basis, detailed data on factors influencing fertility, such as desired family size, birth spacing intentions, and family planning practices, in order to interpret current trends in the birth rate and to prepare more realistic projections of future population growth. In addition, for more effective planning, management, and evaluation of the expanded family planning programs, new information is being provided concerning the social and economic characteristics of couples who are either unable to limit their pregnancies, or who are unable to have all the pregnancies

that they want. The nature and severity of the problems these couples face and the extent to which efforts to help them are succeeding are important aspects of the survey. A major limitation of the survey has been the exclusion from Cycles I and II of single women who do not have offspring living with them. Single women make a significant contribution to the nation's annual fertility, and they will produce approximately one-fourth of total births in the few years immediately following Cycle III of the survey. In expanding the sample for Cycle III to include all women 15-44 years of age, the survey will provide data on the sexually active population as a whole and will fill a significant gap in the need for detailed data on the problems of adolescent pregnancy.

# General population surveys

## National Health Interview Survey

*Background.*—The National Health Interview Survey (NHIS) is a principal source of information on the health of the civilian, noninstitutionalized population of the United States. The National Health Survey Act of 1956 provided for a continuing survey and special studies to secure on a voluntary basis accurate and current statistical information on the amount, distribution, and effects of illness and disability in the United States and the services rendered for or because of such conditions. The survey referred to in the Act, now called the National Health Interview Survey, was initiated in July 1957. In its early years the survey was known to many as the National Health Survey, the name now given to a broader program of surveys in the National Center for Health Statistics.

*Purpose and scope.*—The purpose of the survey is to provide national data on the incidence of acute illness and accidental injuries, the prevalence of chronic conditions and impairments, the extent of disability, the utilization of health care services, and other health-related topics. A major strength of this survey lies in the ability to display these health characteristics by selected demographic and socioeconomic characteristics of the U.S. civilian, noninstitutionalized population.

Because NHIS data are obtained during household interviews from the people themselves, the statistics measure health status and experiences and reflect the social and economic dimensions of health as reported by individuals—the extent and impact of illness and disability and the resulting uses of health care services by the people experiencing them.

Interviews are conducted each week throughout the year in a probability sample of households. The interviewing is performed by a permanent staff of carefully trained and supervised interviewers of the U.S. Bureau of the Census under detailed specifications established by the National Center for Health Statistics. Data collected over the period of a year form the basis for the development of annual estimates of the health characteristics of the population and for the analysis of trends in those characteristics.

The survey covers the noninstitutionalized, civilian population of the United States. Because of technical and logistical problems, several segments of the population are not included in the sample or in the estimates from the survey. Persons *excluded* are patients in long-term care facilities (data are obtained on patients in some of these facilities through the National Nursing Home Survey of the Center); persons on active duty with the Armed Forces (though their dependents are included); U.S. nationals living in foreign countries; and persons who died during the calendar year preceding the interview.

*Sample design.*—The National Health Interview Survey is a cross-sectional household interview survey. The sampling plan follows a multistage probability design that permits the continuous sampling of households. The first stage consists of a sample of 376 primary sampling units drawn from approximately 1,900 geographically defined primary sampling units that cover the 50 States and the District of Columbia. A primary sampling unit consists of a county, a small group of contiguous counties, or a standard metropolitan statistical area. Within primary sampling units, smaller units called segments are defined in such a manner that each segment contains an expected four households. The sampling plan is designed to yield national estimates, but some estimates can be obtained separately for the four geographic regions.

The households selected for interview each week are a probability sample representative of the target population. Each calendar year, data are collected from approximately 40,000 households, including about 110,000 persons. The annual response rate of NHIS is usually at least 95 percent of the eligible households in the sample. The 5-percent nonresponse is divided equally between refusals and households where no eligible respondent could be found at home after repeated calls.

*Data collection procedures.*—Data are collected through a personal household interview conducted by interviewers employed and trained by the U.S. Bureau of the Census according to procedures specified by the National Center for Health Statistics.

All adult members of the household 17 years of age and over who are at home at the time of the interview are invited to participate and to respond for themselves. The mother is usually the respondent for children. For individuals not at home during the interview, information is provided by a responsible adult family member (19 years of age or over) residing in the household. Between 65 and 70 percent of the adults 17 years of age or over are self-respondents. On occasion, a random subsample of adult household members is selected to respond to questions on selected topics. There are also instances in which followup supplements are completed either for the entire household or for individuals identified as having particular health problems. As required, these supplements are either left for the appropriate person to complete and return by mail, or the interviewer calls again in person or by telephone to secure the information directly.

Nationally there are approximately 110 interviewers, trained and directed by health survey supervisors in each of the 12 Census Bureau regional offices. The supervisors are career civil service employees whose primary responsibility is the National Health Interview Survey. The interviewers are part-time employees, selected through examination and testing. Interviewers receive thorough training in basic interviewing procedures and in the concepts and procedures unique to NHIS.

On the average, the interviews require about 45 minutes in the household. Depending upon the family size and the nature and extent of health conditions of family members, the length of interview ranges between 15 and 90 minutes.

*Content of questionnaire.*—The questionnaire consists of two basic parts: a core set of health, socioeconomic, and demographic items; and one or more sets of supplementary health items. The core items constitute approximately 70 percent of the questionnaire and are repeated each year. The arrangement of rotating and single-time supplements allows the survey to respond to changing needs for data and to cover a wide variety of topics, while the core items provide continuous information on fundamental topics. Series 1, No. 11, "Health Interview Survey Procedures, 1957-74," describes changes during that period in survey concepts and questions.

The questionnaire now provides for the following types of core data:

- Basic demographic characteristics of household members, including age, sex, race, education, and family income.
- Disability days, including restricted activity and bed days, and work- and school-loss days occurring during the 2-week period prior to the week of interview.
- Physician and dental visits occurring during the same 2-week period.
- Acute and chronic conditions responsible for these disability days and doctor visits.
- Long-term limitation of activity resulting from chronic disease or impairment and the chronic conditions associated with the disability.
- Short-stay hospitalization data, including the number of persons with hospital episodes during the past year and the number of discharges from short-stay hospitals.
- The intervals since the last doctor and dental visits.

The questionnaire also includes six lists of chronic conditions. Each condition list concentrates on a group of chronic conditions involving a specific system of the body (for example, digestive, skin and musculoskeletal, circulatory, and respiratory). The body systems approach to chronic conditions was adopted because it was found that organizing questions around a particular body system resulted in more thorough reporting and increased the number of conditions for which estimates of prevalence could be made. Beginning in 1978, each of six representative subsamples has been asked questions based on one of the six lists of chronic conditions. In this way national estimates on each of the six body systems are obtained during the same interview year. Prior to 1978 only one condition list was asked each year.

The supplements to the questionnaire change in response to current interest in special health topics. Supplements for 1979 were included on immunization, smoking, home health care, eye care, residential mobility, and retirement income. In 1980 the supplements on home health care, residential mobility, and retirement income were continued and a supplement on health insurance added. Throughout 1981 a child health supplement was included.

Suggestions and requests for special supplements are solicited and received from many sources. These include university-based researchers, administrators of national organizations and programs in the private and public health sectors, and specialists from other parts of the Department of Health and Human Services (for example, the National Institutes of Health and the Centers for Disease Control). Although it is not possible to include all of the suggested topics, every effort is made to be responsive to the data needs of such groups. A leadtime of at least 1 year is required to develop and pretest questions for new topics to be included as special supplements. For example, supplements selected in 1981 will be developed and pretested in 1982 and

implemented in 1983 or 1984. Relevant portions of the questionnaire are included in an appendix to each Series 10 report, and each year's questionnaire is reproduced in its entirety in the annual "Current Estimates" report.

*Data processing and quality control methods.*—Throughout the data collection and processing phases there are extensive quality control activities. Each interviewer edits all completed work before returning it to the regional supervisor, and in the regional office there is a preliminary edit of all questionnaires. As part of the quality control program, interviewers are provided feedback on their errors. In addition, a sample of households is reinterviewed by a supervisor or senior interviewer.

The Census Bureau regional offices forward the questionnaires to the National Center for Health Statistics for coding and data processing. Each questionnaire is checked again at the Center for completeness of field coverage, omissions, and other errors. Illnesses, diseases, and injuries reported by the households are coded to a modified version of the International Classification of Diseases. All coding is subject to recoding on a sample basis to ensure a high level of accuracy.

Thus, potential nonsampling errors such as those in reporting, processing, and nonresponse, which are inherent in any sample survey, are kept to a minimum by methods built into the survey procedures. In addition, with regard to sampling error, standard error charts are created and provided so that a user can calculate the variations in the estimates that might occur because only a sample of the population is surveyed.

*Release of data.*—Data release occurs in several forms, such as publications, tabulations by various sociodemographic characteristics, and public use data tapes. The earliest reports containing National Health Interview Survey data were issued in Series B and C of *Health Statistics from the U.S. National Health Survey*. About 50 reports were published in these series before they were replaced by the *Vital and Health Statistics* series. More than 140 reports from NHIS have been published in Series 10 of *Vital and Health Statistics*. Publication of a year's data begins in about October of the year following completion of data collection and extensive computer editing of the data to ensure quality and reliability.

The first report, "Current Estimates," is followed by 8 to 10 other Series 10 publications. Since not all possible cross-tabulations can be analyzed and published in Series 10 reports, many unpublished tabulations are routinely generated and made available upon request. In addition, these data are published in a number of related special analyses prepared by the Analysis and Epidemiology Program of the Center.

Another medium for data release is public use data tapes. A public use tape is usually available for distribution within about 2 years after completion of data collection. This time lag is due primarily to the need for staff to work with the data in order to ensure completeness and reliability of data obtained from the questionnaire supplements and to permit development of adequate documentation for the tapes. The data tapes, available since the 1969 data collection year, have been widely used by universities and Federal, State, and local governmental agencies.

*Uses of data.*—Generally, uses of NHIS data can be broadly classified into the following areas: program planning and evaluation, public health education and health promotion, epidemiological research, and market research.

NHIS data on morbidity (prevalence estimates) and associated disability (limitation of activity, disability days, and so forth) are used in Federal, State, and local program planning in the assessment of the extent and impact of public health problems. NHIS estimates on trends in incidence, prevalence, and disability are also used extensively in the evaluation of the effectiveness of therapeutic and preventive programs.

Estimates of health problems and resultant access to health care services by various age, sex, occupational, and income subpopulations help identify gaps in the provision and receipt of needed services. Statistics on the prevalence of chronic disease, incidence of injuries, and disability are used in considering the cost of or extending the scope of medical care insurance or compensatory insurance of various types. Estimates of health problems and resultant utilization of health care services were used in formulating the legislation for both the Medicare and Medicaid programs and for making estimates of potential costs of the programs. NHIS data on out-of-pocket health care expenditures and the extent of coverage under private health insurance and public health care programs have been used by the Congressional Budget Office, the Office of Management and Budget, the Office of the Assistant Secretary for Health, and the Social Security Administration to evaluate programs covering the civilian, noninstitutionalized population.

NHIS statistics are used by public, private, and voluntary agencies and organizations for the purposes of public health education and health promotion. For example, NHIS statistics on health characteristics of smokers and nonsmokers appeared in the Surgeon General's reports on Smoking and Health in 1964, 1979, and 1980. In addition, widespread and frequent use of NHIS statistics in newspaper articles, television news spots, popular magazines, and professional journals contributes to the American public's awareness of their health status. This awareness ultimately leads to further individual and organizational health promotion and prevention efforts.

NHIS morbidity statistics are used in the epidemi-

ological analysis of associations between the prevalence of various diseases and sociodemographic characteristics (such as age, sex, and occupation) and in the identification of high-risk subgroups in the population. Identification of geographic patterns in the distribution of diseases can lead to clues for persons researching etiological factors. Researchers also use NHIS data in studies relating to disease to estimate direct costs (for example, expenditures for physician visits and hospitalizations) and indirect costs (for example, economic loss due to premature death, disability, and work loss) due to illness.

Estimates of the number of persons with particular conditions as well as basic sociodemographic attributes are frequently used by private industry not only for estimating potential and current markets for new products or drugs but also for designing advertising campaigns targeted to the affected groups. Specifically, NHIS data on prescription drugs, special aids, corrective lenses, and hearing aids have been used in and of themselves and also in conjunction with the NHIS data on injuries, chronic and acute conditions, and impairments and their impact in terms of limitation of activity and mobility.

About 2,000 data requests are answered annually by the Division of Health Interview Statistics. More than two-fifths of the data requests are from private industry and foundations; almost two-fifths are from Congress and Federal, State, and local government agencies; and the remaining fifth is divided among persons affiliated with educational institutions and private citizens interested in health data.

### National Medical Care Utilization and Expenditure Survey

*Background.*—The National Medical Care Utilization and Expenditure Survey (NMCUES) is a unique source of detailed national estimates on the utilization and expenditures for various types of medical care. This survey builds on the experience of the National Health Interview Survey, the National Medical Care Expenditure Survey, and the former Current Medicare Survey. The first cycle of the survey was conducted to cover calendar year 1980. A second cycle of the survey is planned for the mid-1980's.

*Purpose and scope.*—NMCUES is designed to be directly responsive to the continuing need for statistical information on the health care expenditures associated with health services utilization for the entire U.S. population. Cycle I was designed and conducted in collaboration with the Health Care Financing Administration to provide detailed utilization and expenditure data for persons in the Medicare and Medicaid populations. NMCUES will produce estimates over time for evaluation of the impact of legislation and programs on health status, costs, utilization and illness-related behavior in the medical care delivery system.

Cycle I was composed of several related surveys. The household portion of the survey consisted of a national survey of the civilian noninstitutionalized population and a separate survey of the Medicaid-eligible populations of the States of New York, California, Texas, and Michigan. These two surveys each consisted of five interviews over a period of about 15 months to obtain information on medical care utilization, expenditures, and other health-related information. A third survey, an administrative records survey, was designed to verify the eligibility status of the household survey respondents for the Medicare and Medicaid programs. It also checked insurance claims filed with the national Medicare program and Medicaid programs in each of the four States for persons in the sample of Medicaid eligibles.

*Sample design.*—The national Cycle I household survey comprised persons residing in about 6,000 households. The sample for this survey was a multistage area probability sample drawn from 106 primary sampling units representing the 50 States and the District of Columbia. The State Medicaid household survey sample consisted of about 1,000 families in each of the four States; these families were selected with a known probability of selection from the State Medicaid enrollment lists. Thus, the total sample for the survey was about 10,000 households.

An overall response rate of 89.4 percent was achieved in the first interview for both household surveys in Cycle I: for the national survey the response rate was 91.4 percent, and for the State Medicaid survey the rate was 86.7 percent. Attrition over the course of interviewing resulted in final response rates of 84.9 percent for the National household survey and 76.1 percent for the State Medicaid household survey.

*Data collection procedures.*—NMCUES is a panel study. Interviews in Cycle I were conducted with each household at approximately 3-month intervals, with interviewing beginning in February 1980 and ending in March 1981. The first two interviews were conducted by personal visit of the interviewer to the household, the next two were conducted by telephone (if a telephone was available and acceptable to the household), and the final interview was conducted in person. Each round of interviewing asked about the period following the preceding interview except the first round, which asked about the period of time following January 1, 1980.

Collection of data from the households was facilitated by the use of a calendar and a summary. At the time of the first interview, the household respondent was given a calendar on which to record information about health problems and health services utilization and to assemble physician and other provider bills between interviews. Following each household interview, information about health provider contacts and the payment of charges associated

with them was used to generate a computer summary of information provided. This summary was then printed out in a simple format and mailed to the household for their review for accuracy and completeness prior to the next interview. At the subsequent interview, the interviewers reviewed this information with the household respondent to ensure accuracy and to obtain information not available during a previous interview.

The administrative records survey of Cycle I was a check of the eligibility and claims records of persons reported as covered by Medicare or Medicaid. The supplementary and confirmatory data on Medicaid and Medicare enrollees provided information from the administrative records of the programs for comparison with that reported by the household respondents; the survey was designed specifically to meet the needs of the Health Care Financing Administration for programmatic data that is not otherwise available. For those individuals identified in the surveys as being enrolled in either the Medicare or Medicaid programs, identifying information was collected sufficient to allow the Health Care Financing Administration to flag incoming claims for services paid by Medicare, to allow the State Medicaid agencies in the four States in the State Medicaid survey to abstract data from claims, and to permit all State Medicaid agencies to confirm eligibility and type of eligibility under their State Medicaid program.

*Content of questionnaires.*—Questionnaires for the household surveys were designed to obtain some information on a repeated basis throughout the survey and some information only one time. The repetitive core of questions for Cycle I included health insurance coverage, episodes of illness, the number of bed days, restricted activity days, hospital admissions, physician and dental visits, other medical care encounters, and purchases of prescribed medicines. For each contact with the medical care system, data were obtained on the nature of the health conditions, characteristics of the provider, services provided, charges, sources, and amounts of payment. Questions asked only once included data on access to medical care services, limitation of activities, occupation, income, and other sociodemographic characteristics.

*Data processing and quality control methods.*—Data collection and processing phases incorporated a number of quality control activities, including a detailed edit by each field interviewer and, immediately upon receipt at the central office, a premachine edit of all key linkage variables. Errors discovered in this premachine edit and in further edit of a sample of all items provided feedback to interviewers and supervisors.

Illness conditions, operations performed, specialties of medical providers not precoded on the questionnaire, and the industries in which persons were employed were coded prior to data entry. Quality control samples provided continuous monitoring of each coder, with retraining of interviewers and recoding of data, as appropriate, to ensure high levels of accuracy. Data entry through key to disk minicomputers provided a further level of editing for valid ranges and codes. Check digits embedded in key identifying numbers kept miskeyed linkage numbers to an absolute minimum.

*Release of data.*—Data from Cycle I of NMCUES will be released in several forms. The publications program is a collaboration between the National Center for Health Statistics and the Health Care Financing Administration. Preliminary data will be published in *Preliminary Data* reports, similar to the Center *Advance Data* reports. More detailed series reports based on the survey findings will be issued in a joint series with the Center preparing reports pertaining to the population as a whole and the Health Care Financing Administration preparing reports specific to the Medicare and Medicaid populations. A public use data tape is planned with release targeted for late 1982. The tape will actually be several sets of files including the national household data and the State data for each of the four States in the administrative records survey. Each set of files will contain data on the family, the individual, and the health care visit, pertaining to expenditures, utilization of services, and measures of health. In addition, unpublished tabulations will be made available; and, as resources, permit, special tabulations will be prepared upon request.

*Uses of data.*—The National Center for Health Statistics has the responsibility for meeting the data needs of the Department and other users for expenditure and utilization data. The NMCUES data provide a statistical base for major health policy decisions on key issues such as health care financing and cost containment. Data are available for the Medicare and Medicaid populations and non-Federal beneficiaries so that comparisons can be made. Also, some data were collected on the State level for Medicaid beneficiaries because of the high degree of variability of eligibility requirements and service coverage in the State-managed Medicaid programs.

In 1979 national health care expenditures amounted to $212.2 billion, about 9 percent of the gross national product; and the figures have been growing steadily for several decades. Of this $212.2 billion, 43.1 percent was paid from public sources in 1979. These figures, compiled by the Health Care Financing Administration, are aggregates. NMCUES is designed to provide statistics on the effects of health care utilization and expenditures at the individual and family level. In gross terms, who is using what types of medical care, how much is it costing them, and how is the bill being paid?

NMCUES provides a unique source of extremely

detailed health care utilization and financing information, which will enable detailed measurement of utilization and expenditures for different subgroups of the population. Specific items of service are identified that can be related to chronic and acute illness, disability, work-loss days, and limitation of activity. Such data are essential to the measurement of the impact of legislation and health insurance programs on the charges for and financing of health services and insurance. This survey will be the principal source of ongoing data needed to monitor changes over time in the impact on individuals and families of private and public programs. Specific uses of the data include the following:

- Providing a statistical base for the health care cost monitoring efforts of the Department of Health and Human Services.
- Providing updated, comparable measures of utilization and expenditures for the design and monitoring of health insurance proposals.
- Providing data on trends and costs of health care services for different population subgroups over time.

### National Health and Nutrition Examination Survey

*Background.*—The first National Health and Nutrition Examination Survey, referred to as NHANES I, was initiated in 1970, with data collection beginning in April 1971. NHANES I was a modification and expansion of the Health Examination Survey (HES), which had been initiated a decade earlier and had carried out three separate programs. The restructuring and modification of HES reflected the assignment to the National Center for Health Statistics of an additional specific responsibility—the measurement of the nutritional status of the population and the subsequent monitoring of changes in that status over time. Upon completion of NHANES I, data were collected for a second National Health and Nutrition Examination Survey (NHANES II) beginning in February 1976 and finishing in February 1980. A special health and nutrition examination survey directed to persons in families having one or more members of Hispanic origin or descent who live in areas with a high concentration of Hispanics has been planned for the period between completion of NHANES II and the beginning of NHANES III (now planned to begin data collection in 1987). Pretesting for the Hispanic HANES began in 1981, and data collection will begin in 1982.

*Purpose and scope.*—NHANES and its predecessor, HES, share a common purpose—the collection and utilization of data that can be obtained best or only by direct physical examination, clinical and laboratory tests, and related measurement procedures. This information, which cannot be furnished by the people themselves or by the health professionals who provide their medical care, is of two kinds. Prevalence data are collected for specifically defined diseases or conditions of ill health; and normative health-related measurement data are collected that show distributions of the total population with respect to particular parameters such as blood pressure, visual acuity, or serum cholesterol level.

The surveys use probability samples of the U.S. population to provide representative national data that are analyzed and made available in a series of reports. Successive surveys in the HES and NHANES programs have been directed to different segments of the population and have had different sets of target conditions. Thus, the first Health Examination Survey, or "cycle," involved examining a sample of adults with the focus primarily on selected chronic diseases. The second and third cycles of HES were directed respectively to children between the ages of 6 and 11and to youths between the ages of 12 and 17. Both of these surveys emphasized growth and development data and sensory defects. The nutrition component of the first NHANES program was directed to a probability sample of people in the broad age range 1 to 74 years, while the detailed health examination component focused on the population between ages 25 and 74. NHANES II was again directed to a broad population aged 6 months to 74 years, and the data on nutrition that were collected will be used in conjunction with the earlier NHANES I data to monitor changes in nutritional status over time.

*Sample design.*—The samples for all of the HES and NHANES programs have been multistage, highly clustered probability samples. All of the samples were stratified by broad geographical region and by population density grouping. Within the strata, the sampling stages employed have been the primary sampling unit, the census enumeration district, the segment, the household, and lastly the individual person. Until the household stage was reached, all sampling was carried out centrally in conjunction with the U.S. Bureau of the Census.

The next stage of the sampling was conducted in the field in the particular chosen area. It involved interviewer visits and questionnaire completion at each one of the selected households, with the final selection of individuals included in the sample being dependent upon information elicited by the household interview questionnaire. The size of the sample in the survey program has varied. In each of the three HES programs, the sample size was approximately 7,500 persons. In NHANES I the sample selected for the major nutrition components of the examination contained approximately 28,000 people and yielded about 21,000 examined persons. A sample compara-

ble in size for NHANES II again yielded approximately 21,000 examined persons.

*Data collection procedures.*—Household interviewing during NHANES II was conducted by Bureau of the Census personnel. This represented a departure from previous surveys in which Center employees served this function. However, Center employees did the rest of the interviewing, as well as the history taking, examining, testing, and measuring in the mobile examination centers. Data collection teams consisted of specially trained interviewers and examiners, including physicians, nurses, dentists, dietitians, and medical, laboratory, and X-ray technicians.

The examinations take place in the survey's specially constructed mobile examination centers, each consisting of three truck-drawn trailers. These trailers are interconnected and provide a standardized environment with equipment for the performance of specific parts of the examination. This standardized environment is necessary for such components of the examination as audiometry, which requires hearing chambers within which the ambient noise level conforms to the American Speech Association standards for acoustical measurements.

The general pattern of data collection has meant that each survey has been conducted over a period of 3 or 4 years. This is due to the limitations on the number of persons examined in a given time span (for example, the number of field teams and the number of sample areas). The kinds of data to be collected are also limited, since conditions that might show marked year-to-year variation or seasonal patterns cannot be included. However, many important chronic diseases and health-related measurements are not subject to such changes in prevalence within shortrun periods. The distribution of the population according to unassisted visual acuity levels and the prevalence of such conditions as diabetes or hypertensive heart disease may vary over long periods of time, but not so rapidly as to prevent data collection over a 3- or 4-year period from giving a correct picture of the population levels during or at the midpoint of that period.

Voluntary sample surveys present a problem if no data are collected on a large fraction of the selected sample because individuals selected to participate in the program are not willing to do so. In the HES and NHANES program there has always been and continues to be much attention devoted to the question of the response rate, the proportion of sample persons who are actually examined. In NHANES there have been, as anticipated, more problems in the area of response than had been encountered in the earlier HES programs. The difficulties faced have led to a variety of innovative measures, including a policy of remunerating examined persons. As a result, both NHANES programs succeeded in obtaining household interview data on about 99 percent of the sample population. More detailed health data appear in the medical history questionnaires; these were completed for 88 percent of the selected sample persons in NHANES I and for 91 percent of the selected sample persons in NHANES II. Finally, in NHANES I, 74 percent of the sample persons selected for the nutrition component and 70 percent of the persons selected for the detailed health component were given the standard examinations and tests; in NHANES II the overall response rate for the examination component was 73 percent. There is considerable ancillary information on most of the persons in the sample population who were not examined, and it is possible to make use of that data in the process of imputation and analysis of nonresponse bias. There is, moreover, some evidence that data obtained through examinations, tests, and measurements such as those used in these surveys are less susceptible to potential bias from a given rate of nonresponse than are data provided by the individuals themselves.

*Content of questionnaires, tests, and examinations.*—The kinds of information collected in the NHANES and other examination survey programs are so varied and extensive that they are only illustrated here. With respect to nutrition, four types of data are included:

- Information concerning dietary intake—taken from 24-hour recall interviews and food frequency questionnaires, both administered by an interviewer who has been a trained dietitian.
- Hematological and biochemical tests—a sizable battery of such tests, with processing at the mobile examination centers where necessary, but for the most part at a central nutrition laboratory established at the Centers for Disease Control.
- Body measurements—an especially important battery in connection with infants, children, and youths, where growth may be affected by nutritional deficiencies.
- Various signs of high risk of nutritional deficiency—based on clinical examinations.

The health (as distinguished from nutrition) component of the NHANES program includes detailed examinations, tests, and questionnaires, which have been developed to obtain a measure of prevalence levels of specific diseases and conditions. These vary with the particular program and have included such conditions as chronic rheumatoid arthritis and hypertensive heart disease. Important normative health-related measurements, such as height, weight, and blood pressure, are also obtained.

A major element in the health component of NHANES I was an assessment, using index conditions, of unmet health needs. For the index conditions the examination established whether or not emphysema or another chronic respiratory disease

was present. At the same time, information was obtained from the examined person with respect to self-perceived health needs and actions taken in seeking medical care. Analytic plans call for the interrelating of these two kinds of information to produce measures of unmet health care needs.

In NHANES II the nutrition component remained nearly identical to that fielded previously. From an early analysis of the NHANES I data, it was decided to focus the nutritional examination elements on an anemia-related assessment approach. This involved the addition of certain medical history items and a more tailored set of laboratory determinations. Less was done in the area of health care needs. However, the emphasis placed on the effects of the environment upon health will probably be continued in future programs. Data were gathered to measure the levels of pesticide exposure, the presence of certain trace elements in the blood, and the amounts of carbon monoxide present in the blood. In the medical area, primary emphasis in NHANES II was placed on diabetes, kidney and liver functions, allergy, and speech pathology.

The time required for the examination varies with the content of the examination and the age of the examinee. The time constraint included among planning factors has been that the total examination time not exceed 2-2½ hours. Much attention is given in the planning process and in the actual pretests of a survey to the actual flow of examinees through the examination center, and every effort is made to streamline this process in order to reduce the time burden on the sample person. Additional respondent burden arises from the interview, from the completion of forms and questionnaires in the household, and from the varying time required to travel to the examination site. Data collection forms used in HES I, II, and III, NHANES I, the NHANES I Augmentation Survey, and NHANES II are reproduced in Series 1, reports numbered 4, 5, 8, 10b, 14, and 15, respectively.

*Data processing methods.*—The data collected in NHANES require a variety of data processing methods. X-rays must be interpreted; blood and other laboratory specimens must be processed through the appropriate laboratory operations. Certain data, such as electrocardiographic tracings, record directly onto magnetic tape and must undergo appropriate processing to be translated into digital tape form and subsequently to be interpreted. Examination record forms and interview and questionnaire data must be coded and put onto magnetic tape.

Methods for handling specific examination elements are chosen on the basis of appropriateness. The program makes use of precoded forms, marked-sense record forms, self-administered forms, interviewer-administered forms, automatic recording devices, photographs, and so forth. Throughout whatever process is chosen, constant emphasis is placed on quality control measures such as editing, verifying, and replicating. Finally, outside consultants frequently collaborate in the analysis and reporting of the resultant data.

*Release of data.*—Findings from NHANES programs traditionally are presented primarily through publication of individual reports in the National Center for Health Statistics *Vital and Health Statistics*, Series 11. In order to make NHANES findings more rapidly available to consumers, a program has been initiated to release basic descriptive summary tables in what are called "basic data publications." These publications are an adjunct to the other Series 11 reports for selected topics. Additionally, some NHANES data appear in abbreviated preliminary reports from *Advance Data.*

Data are also released through appropriate scientific journals, separate monographs, special reports, and data tapes. Because of the voluminous amount of data collected in NHANES I and the need for disseminating data as expediently as possible, a policy shift was made that allows for the simultaneous release of data tapes to in-house analysts and sale of these tapes to the public as soon as final editing has been performed and documentation prepared. Data tapes from NHANES II are also being released upon completion of final editing and preparation of documentation.

Published reports are not issued on a set frequency but rather made available as completed. The reports are organized on a topical basis with, for example, one report presenting data on periodontal disease, another on auditory acuity, and so forth. Generally, the first reports produced are descriptive in nature while later reports become more analytic and may be interdisciplinary in approach.

*Uses of data.*—The diverse nature of the data attracts a wide variety of users. In addition to policy makers and planners in Federal, State, and local health agencies, users of NHANES data include large numbers of health workers in the fields of nutrition, child health, pediatrics, anthropology, medical education, and health research, who have requested reports on body measurements, growth, and development. Data on body measurement have also been used by many industrial groups, including automobile manufacturers. Since some of the primary users of NHANES and HES data have been authors of text and reference materials used in medical schools and by health workers, these normative data reach a vast audience.

The sale of data tapes from this program has increased greatly during the past few years, with known data users numbering in the hundreds. Several colleges and universities have purchased data tapes to be included in their computer tape libraries. Some of the users are medical research institutions and individ-

ual researchers who wish to explore in detail specific elements of the data. Projects currently underway using NHANES data tapes include research on the effects of alcohol intake on health, research on the consumption of food items for marketing studies, small scale or local health surveys that use the NHANES data as norms, and epidemiological studies that examine geographic patterns in morbidity.

## National Hospital Discharge Survey

*Background.*—The National Hospital Discharge Survey (NHDS) is the principal source of information on inpatient utilization of short-stay hospitals. Data collection began in 1964 and has been continuous since then.

*Purpose and scope.*—The purpose of NHDS is to produce statistics that are representative of the experience of the U.S. civilian population discharged from short-stay hospitals. Specifically, the survey provides information on characteristics of patients, lengths of stay, diagnoses, surgical operations, and patterns of use of care in hospitals of different size and ownership in the four major geographic regions of the country.

The scope of NHDS is limited to discharges from non-Federal hospitals in the 50 States and the District of Columbia. Only hospitals with six or more beds and an average length of stay for all patients of less than 30 days are included in the sample.

*Sample design.*—The unit of enumeration in the survey is a hospital discharge. The sample plan is basically a two-stage stratified design. The first stage is a sample of about 10 percent of the short-stay hospitals, excluding Federal hospitals, listed in the National Master Facility Inventory (the sampling frame—see section "National Master Facility Inventory"). The primary stratification variables are bed size and geographic region. Hospitals are selected in direct proportion to size so that hospitals with 1,000 or more beds are selected with certainty, and hospitals with less than 50 beds are sampled with a probability of approximately 1 in 40. Growth in the inventory of hospitals is represented in the survey by a sample of hospitals selected from a special universe of new hospitals.

The second stage of the design is a systematic sample of the discharges from the sampled hospitals. The sampling frame in nearly all hospitals is the daily listing of discharges. The size of the within-hospital sample varies inversely with the size of the hospital, from about 1 per 100 discharges in hospitals with 1,000 or more beds to 4 per 10 discharges in hospitals with less than 50 beds. The overall sampling rate for each bed-size group is about 1 per 100 discharges, the product of the first- and second-stage sampling rates.

In 1979 the sample consisted of 544 hospitals from a universe of approximately 8,000 short-stay hospitals. Of the 496 hospitals in the scope of the survey, information was collected from 416 participating hospitals (an approximately 84-percent response rate) on approximately 215,000 discharges.

*Data collection procedures.*—The U.S. Bureau of the Census, acting as the data-collecting agent for the National Center for Health Statistics, inducts sample hospitals into NHDS. After induction, hospitals are visited at least once a year by a representative of the Bureau of the Census. At that time survey procedures are reviewed, and information about the hospital is updated.

Discharge data are collected throughout the year. Sample discharges are systematically selected, usually on the basis of the final digit or digits of the patient's medical record number. For each sample discharge, an abstractor records personal, administrative, diagnostic, and surgical information from the face sheet of the patient's medical record onto a medical abstract form. Data collection frequency depends on the arrangement made with the hospital. In about 35 percent of the participating hospitals, a representative of the Bureau of the Census visits the hospital every 2 months, completes the abstract forms for records selected during the previous visit, and selects records for abstracting at the next visit. This allows time for records to be completed and properly filed (or pulled from file) prior to the visit. In about 65 percent of the hospitals, the same forms are completed by members of the hospital medical record department. All completed forms are forwarded to one of the Census regional offices for review and then to the National

Center for Health Statistics for coding and data processing operations.

*Content of medical abstract form.*—The medical abstract form contains items relating to the personal characteristics of the patient, including date of birth, sex, race, and marital status, but not name and address; administrative information, including admission and discharge dates, discharge status, and medical record number; and medical information, including diagnoses and surgical operations or procedures. It is estimated that medical record personnel can sample and complete each form, on the average, in about 5 minutes.

The contents of the medical abstract form did not change from the inception of the survey until 1977, when modifications were made so that it more nearly parallels the Uniform Hospital Discharge Data Set. The items added to the abstract at that time are residence of patient (zip code), expected source of payment, disposition of patient, and dates of procedures. In 1968-70 actual hospital charges by service and payments by source were recorded on a ledger abstract form for approximately one-third of the sample discharges. A copy of the medical abstract form is included as an appendix in Series 13 reports based on data from NHDS.

*Data processing methods.*—After transmittal from the hospital, abstract forms are subject to two reviews, two machine edits, and two quality control procedures. Forms are reviewed for completeness at the Census regional offices and either forwarded to the National Center for Health Statistics or returned to the hospital for more information. On receipt at the Center forms are again checked and, if necessary, returned to the hospital. After review, up to seven diagnoses and four operations are coded according to the clinical modification of the International Classification of Diseases—ICD-9-CM. Coded information is then keyed from abstract to disk, and a preliminary machine-edit program checks for missing, invalid, and inconsistent codes. The information is corrected, if necessary, and then transferred from disk to computer tape. For 1970-78 data, up to five diagnoses and three operations were coded according to the eighth revision of the International Classification of Diseases—ICDA-8.

*Release of data.*—Annual data are published in *Advance Data* and in the Center's *Vital and Health Statistics* Series 13. The publication program is, at a minimum, to update nonmedical, medical, and surgical data for characteristics of patients and hospitals. Special reports on average length of stay, patient charges, geographic utilization, and methodology are also published.

Unpublished data are available on request from the Hospital Care Statistics Branch, which receives about 1,000 requests per year, usually for specific diagnostic and surgical listings in the International Classification of Diseases. In addition, data for the years from 1970 to the present are available on public use data tapes.

The names of the participating hospitals and all information related to individual patients are confidential. No data are released in published, unpublished, or tape form that could identify hospitals or patients.

*Uses of data.*—The National Hospital Discharge Survey is the principal source for national data on the characteristics of patients discharged from short-stay hospitals. The data are used for a variety of planning, administrative, and evaluative activities by governmental, professional, scientific, educational, and commercial institutions, as well as by private citizens. The users include Federal agencies such as the National Institutes of Health, the Centers for Disease Control, and the Indian Health Service; universities and medical schools; professional organizations such as the American Medical Association, American Hospital Association, and World Health Organization; hospitals; medical research laboratories; pharmaceutical and medical supply manufacturers; publishing houses; market research groups; and insurance companies.

### National Ambulatory Medical Care Survey

*Background.*—In May 1973 the National Center for Health Statistics inaugurated the National Ambulatory Medical Care Survey (NAMCS) on a continuing basis to gather and disseminate statistical data about ambulatory medical care provided by office-based physicians to the population of the United States. The need for such a system had been recognized many years before, having been pointed out as early as 1953 by the Subcommittee on National Morbidity Surveys of the U.S. National Committee on Vital and Health Statistics. During a 5-year feasibility study period beginning in 1967 three major national pilot studies and several small area studies, combined with extensive consultation with experts in various areas of medical care delivery, were employed in developing the current NAMCS instruments and procedures. Through this process the current NAMCS materials and procedures have been refined to request only information considered essential to describe adequately the utilization of ambulatory services and to require a minimal amount of time for the participating physicians.

*Purpose and scope.*—The purpose of NAMCS is to meet the needs and demands for statistical information about the provision of ambulatory medical care services in the United States. Ambulatory services are rendered in a wide variety of settings, including physicians' offices, neighborhood health centers, and hospital outpatient facilities. It is expected that NAMCS will eventually encompass all of these settings and

that appropriate survey instruments and methodology will be developed as resources permit. Currently, the NAMCS target population consists of all office visits within the conterminous United States made by ambulatory patients to non-Federal physicians who are in office-based practice and engaged in direct patient care. Excluded are visits to hospital-based physicians; visits to specialists in anesthesiology, pathology, and radiology; and visits to physicians who are principally engaged in teaching, research, or administration. Telephone contacts and visits other than office visits are also excluded. Since about 70 percent of all direct ambulatory medical care visits occur in physicians' offices, the current NAMCS design provides data on the majority of ambulatory care services.

*Sample design.*—The most objective and reliable sources of data about physicians' services rendered to ambulatory patients during office visits are the physicians themselves and members of their office staff. The sampling frame is a list of licensed physicians in "office-based, patient care" practice compiled from files that are classified and maintained by the American Medical Association and the American Osteopathic Association. These files are continuously updated by both associations, making them as current and correct as possible at the time of sample selection.

NAMCS uses a modified probability-proportional-to-size sampling procedure using separate sampling frames for standard metropolitan statistical areas and for nonmetropolitan counties. After sorting and stratifying by size, region, and demographic characteristics, each frame is divided into sequential zones of 1 million residents, and a random number is drawn to determine which primary sampling unit from each zone is included in the sample. The NAMCS final first-stage sample contains 87 primary sampling units, corresponding to individual counties or small groups of contiguous counties across the country.

The second-stage sample is selected from the list of physicians located in the sample primary sampling units, ordered by major specialty categories so that the overall probability for including any individual physician is approximately constant. The present annual sample consists of approximately 3,000 physicians. The sample physicians are randomly distributed across the 52 weeks of the year so that the resulting data reflect any seasonal variations. Since the assignment of the reporting week is an integral part of the sample design, each physician is required to report during his predetermined period, and no substitute reporting periods are permitted. Approximately 75 percent of the eligible physicians in the sample participate in the survey. From this size physician sample, information is secured from about 50,000 patient visits a year. Samples for subsequent years exclude, with certainty, sample physicians included in the previous 2 years of the study.

The final stage involves sampling patient visits within a physician's practice. The sampling rate, which is determined at the time of the interviewer's appointment, is dependent on the number of days during the reporting week that the physician is in practice and the number of patients expected to be seen. In actual practice, the sampling procedure is handled through the use of a Patient Log (see "data collection procedures").

*Field procedures.*—To maximize participation levels and minimize data reporting burden in the physician's office, NAMCS field procedures have been designed to accommodate the circumstances of individual physicians. Each physician is contacted by several means, including mail, telephone, and personal interview. Initially, each physician in the sample is sent an introductory letter from the Director of the Center. The physician is then telephoned by an informed and trained interviewer who explains the survey briefly and arranges a personal appointment to relate more detailed instructions. During this appointment, the interviewer verifies the physician's eligibility for participation in the survey, delivers survey materials with printed instructions, provides detailed verbal instructions, and assigns a predetermined 7-day (Monday through Sunday) reporting period. Also, during this appointment data concerning basic practice characteristics, such as the physician's specialty and staff makeup, are collected for use in analysis.

At the beginning and again during the physician's assigned reporting week, the interviewer contacts the sample physician to answer possible questions and to ensure that procedures are going smoothly. At the end of the week, the participating physician mails the finished forms to the interviewer, who edits them for completeness before transmitting them for central data processing.

*Data collection procedures.*—The actual data collection for the National Ambulatory Medical Care Survey is carried out by the participating physician, aided by office assistants when possible. The physician completes a Patient Record for a sample of his patients seen during his assigned reporting week. Based on the physician's own estimate of patients expected to visit during the survey period, the physician is assigned to use an "every-patient" or a "patient-sampling" procedure. These sampling procedures are designed so that Patient Records are completed each day of practice for, at most, 10 patient visits. Physicians expecting 10 or fewer visits per day record data for all of them, while those expecting more than 10 visits per day record data after every second, third, or fifth visit, using a random start and observing the same predetermined sampling interval continuously. These procedures minimize the workload of data collection and maintain equal reporting levels among sample physicians, regardless of the size of their practice. Each form requires 1 to 2 minutes to

complete, so that approximately 15 minutes are required on days when patients are attended in the physician's office.

Two data collection forms are employed by the participating physician: the Patient Log and the Patient Record. The Patient Log is a sequential listing of patients that serves as a sampling frame to indicate which visits data should be recorded for. The Patient Record varies slightly from year to year, but it generally contains items of information about the visit such as its date and duration; the patient's date of birth, sex, race, ethnicity, and reason for visit; whether the patient had been seen for the particular problem before and whether the patient was referred by another physician; the length of time since onset of the problem; diagnoses; diagnostic and therapeutic services; and disposition. Periodically, supplementary items are added to the basic Patient Record to investigate specific health conditions or other aspects of ambulatory care. For example, questions about specific medications were added for 1980-81, and questions about accident- or product-related illnesses were added for 1979. A copy of the Patient Record is included in the annual "Summary: NAMCS" *Advance Data* report and as an appendix in Series 13 reports based on NAMCS data.

*Data processing methods.*—Initial edits for completeness of the Patient Record are done by the interviewer as the forms are received from the participating physician. These forms are then sent to central processing, where they are clerically edited for consistency. The two medical questions are then coded: the patient's reasons for visit are coded using the NAMCS "Reason for Visit Classification" (described in Series 2, No. 78) and the physician's diagnoses are coded using the International Classification of Diseases. The data are then keyed onto magnetic tape, at which time additional machine edits and consistency checks are run.

*Release of data.*—NAMCS results in the form of summary statistical tabulations of national and regional estimates for number of visits, percent distributions, and population rates of use are published as soon as possible after each annual cycle is complete. More detailed tabulations and analyses follow, which present visit characteristics by major physician specialty groups, patient groups, diagnostic categories, treatment provided, and disposition. Cross-tabulations of less common visit characteristics are published when sufficient data about them are available to meet practical standards of precision. These data are released through the Center's *Advance Data* and *Vital and Health Statistics* Series 13 reports. Other publications, including the various journals and newsletters of the NAMCS endorsing organizations, also provide a means of releasing data to the medical community. In addition, public use data tapes are made available approximately 1 year after the end of data collection, when documentation of the data is complete and its validity is checked.

As part of the procedures designed to protect the identity of the patients, the Patient Logs containing the names of the patients are detachable from the Patient Records. At the end of the reporting period the Patient Logs, and thus the names of the patients, are retained by the physician. Sample physicians are further assured that all information that would permit identification of an individual, a practice, or an establishment will be held confidential. All findings are released in the form of summary statistics that preclude any individual identification, and all identifying data are destroyed as soon as data processing operations are complete.

*Uses of data.*—NAMCS provides a wide range of data on the characteristics of the users and providers of ambulatory medical care. These baseline data, together with trend data as the survey progresses, provide new insights into ambulatory medical care and stimulate further research on the use, organization, and delivery of ambulatory care.

The manner in which ambulatory care is actually dispensed, as described by the survey, can be related to arrangements for educating and training physicians and other health personnel to provide ambulatory medical care, particularly primary care. Such comparisons suggest useful new directions and emphases that can be incorporated in the curricula, helping to make medical education more responsive to the shifting requirements and health problems of the population. NAMCS information is useful to health planning agencies; managers of health delivery systems; and others concerned with planning, monitoring, and managing health care resources. It is valuable to those who develop and evaluate new and modified health care systems and arrangements. The continuing nature of the survey permits observation and measurement over time of different modes for managing and treating patients' problems, and it provides general information on the etiology and epidemiology of selected conditions. It also provides valuable information about the speed and effectiveness with which certain advances in medical practice are adopted by the basic source of medical care, the office-based physician.

### National Nursing Home Survey

*Background.*—Between 1963 and 1969 the National Center for Health Statistics conducted surveys of nursing homes and their residents on an *ad hoc* basis. With the implementation of the Medicaid and Medicare programs, the increased utilization of nursing homes, and the projected increases in the aged population, those who set standards for, plan, provide, and assess long-term care services needed comprehensive national data on a continuing basis. To meet their

needs, the National Nursing Home Survey was developed in 1972, with the initial survey conducted in 1973-74 and the second conducted in 1977.

*Purpose and scope.*—This periodic data collection system is a series of nationwide sample surveys of nursing homes, their residents, and staff. The purposes of the surveys are as follows:

- To collect national baseline data on characteristics of the nursing home, its services, residents, and staff for *all* nursing homes in the Nation, regardless of whether or not they are participating in Federal programs such as Medicare or Medicaid.
- To collect data on the costs incurred by the facility for providing care by the major groupings—labor, fixed, operating, and miscellaneous costs.
- To collect data on certification for participation in the Medicare and Medicaid programs (such as the utilization of certified beds and the health of residents receiving program benefits) so that all data can be analyzed by certification status.
- To provide comparable data for valid trend analyses on a variety of topics. Such analyses can, for example, identify the impact of legislative changes in standards and in reimbursement on the growth of facilities, and of the impact of institutionalization on the health of the aged.
- To interrelate facility, staff, and resident data to reveal the relationships that exist between utilization, services offered, charges for care, and the cost of providing care.

For the initial survey conducted in 1973-74, the universe included only those nursing homes that provided *some level of nursing care*, regardless of whether or not they were participating in the Medicare or Medicaid programs. Thus, homes providing only personal or domiciliary care were excluded. Beginning with the 1977 survey, the universe was expanded to include *all* nursing, personal care, and domiciliary care homes, regardless of their participation in Medicare or Medicaid. Homes that provide room and board only are excluded. In both surveys, homes in the universe included those that were operated under proprietary, nonprofit, and government auspices. The universe also included homes that were units of a larger institution (usually a hospital or retirement center).

*Sample design.*—The National Master Facility Inventory (NMFI) listing is the universe from which the sample homes are selected. The NMFI listing, maintained by the Center, contains the basic information about each home (such as name, address, size, ownership, number of residents, and number of staff) that is needed to design efficient sampling plans.

Resident data are collected by reviewing medical records and questioning the nurse who usually provides care for the resident. Residents are not interviewed directly. The initial National Nursing Home Survey, conducted from August 1973 to April 1974, had a nationally representative sample of 2,100 nursing homes, with a subsample of 25,100 staff and 19,400 residents. The second survey, conducted from May to December 1977, had a total sample of 1,700 nursing homes, with a subsample of 16,800 staff, 7,100 residents, and 5,300 discharged residents.

*Data collection procedures and questionnaire content.*—The survey system uses several questionnaires. The facility questionnaire includes questions on number of beds and residents, services provided, certification status, and various utilization measures. The expense questionnaire includes questions on the facility's expenses by the major groupings—labor, fixed, operating, and miscellaneous. The staff questionnaire includes questions about the employee's demographic characteristics, work experience, education, and salary. The residents questionnaire includes questions about the resident's demographic characteristics, health status, functional status, participation in social activities, monthly charge, and source of payment. Included in the 1977 survey was a questionnaire for discharged residents. This included some of the same items as the questionnaire for current residents, items selected on the basis of their availability in the medical record.

The survey has a number of respondents in each nursing home and is a combination of personal interview and self-administered questionnaires. Facility information is secured through a 20-minute personal interview with the administrator. Expense data are collected on a self-administered questionnaire, requiring about 30 minutes to answer, completed either by the facility's accountant under authorization from the administrator or by the administrator. Sampled staff members complete a brief form that requires about 5 minutes to complete. Information on sample current residents is secured by the interviewer in a personal interview with the nurse who provides care to the resident and who refers to information from the medical record. About 15 minutes is required for each sample resident. For the 1977 survey, information on the sample discharged residents was secured by the interviewer in a personal interview with the nurse who was most familiar with the medical records and who referred to them for replying to all questions.

All survey questionnaires for the 1973-74 survey are included in Series 13, No. 22; and the questionnaires used in the 1977 survey are contained in Series 13, No. 43. Selected questionnaires from the surveys are found in other Series 13 reports based on the National Nursing Home Survey.

The 1977 survey included several modifications of the 1973-74 survey design and methodology.

- It collected data on discharged residents, especially in the areas of health status, length of stay,

discharge status, and where the resident was to go after discharge.

- It collected data on the revenues of the facility as well as on its expenses.
- It produced estimates for the five States with the largest proportion of nursing home residents.
- It limited respondent reporting by limiting subsamples of residents and discharges to a maximum of 8 per facility, and of staff to 23 per facility.

The response rates for the surveys differed according to the type of questionnaire used, as presented in table B.

*Data processing methods.*—All data items are edited to ensure that all responses are accurate, logical, and complete. Those items left unanswered are imputed by assigning a value from a responding unit with major characteristics identical to those of the nonresponding unit.

*Release of data.*—National estimates are available to requesters in the form of tabulations, Center publications, and data tapes. Tables are aggregated so that no one respondent can be identified. Similarly, no identifying information concerning the facility, staff, or residents is contained on the data tapes, although a pseudoidentification number allows linkage of data among the various files.

A report analyzing provisional tabulations is available 6 months after data collection ends. Selected final tabulations are released 6 months later. Data are analyzed and presented in the Center *Advance Data* reports and in *Vital and Health Statistics* Series 13 and 14. These reports present data on facilities (utilization, expenses, and services), residents (demographic characteristics, health status, services received, and charges), discharged residents, and staff (number, training, and experience). Once data have been reviewed and analyzed, data tapes are available for release to requesters.

Table B. Response rates for National Nursing Home Surveys, by type of questionnaire used (percent): United States, 1973-74 and 1977

| Questionnaire | Response rate | |
|---|---|---|
| | 1977 survey | 1973-74 survey |
| Facility | 95 | 97 |
| Expense | 85 | 88 |
| Staff | 81 | 82 |
| Residents | 99 | 98 |
| Discharges | 97 | ... |

*Uses of data.*—Nursing homes are the most rapidly expanding sector of the health care industry. Federal and State funds spent for nursing home care have been increasing each year. Thus, national data on the costs and characteristics of the nursing home, its residents, discharged residents, and staff, are needed on a continuing basis by the Department of Health and Human Services for program planning and the setting of national policies. Congress needs information on the interrelationships among facility-resident-cost data in order to draft legislation on standards and reimbursement. In addition, these data are needed by professionals who conduct epidemiological research and those who plan, provide, and assess long-term care services at State and local levels.

### National Master Facility Inventory

*Background.*—The National Master Facility Inventory (NMFI) is a comprehensive file of the 33,000 facilities in the United States that provide medical, nursing, personal, or custodial care to groups of unrelated persons on an inpatient (at least overnight) basis. Facilities in NMFI are categorized into three broad types: hospitals, both short- and long-stay; nursing and related care homes; and other custodial or remedial care facilities, including homes or resident schools for the deaf, blind, mentally retarded, emotionally disturbed, and other neurologically impaired or physically handicapped; resident treatment centers for alcohol and drug abusers; orphanages or homes for dependent children; and homes for unwed mothers. It is the most comprehensive file of inpatient health facilities available in the United States.

*Purpose and scope.*—NMFI has two basic purposes. It is an important national source of statistics on the number, type, and geographic distribution of inpatient facilities in the United States. In addition, it serves as the universe from which probability samples are selected for conducting sample surveys.

*Data collection procedures.*—NMFI was first assembled in 1962-63 by collating the files of four Federal agencies that contained the names and addresses of facilities, the directories of national associations and organizations, and State licensure files.

Two mechanisms are used to keep the data in NMFI as current as possible. For all facilities except hospitals, the National Center for Health Statistics conducts a series of mail surveys, first, to ensure that the data on file on the basic characteristics of the facilities are accurate and, second, to identify and then delete those facilities that have gone out of business or are no longer eligible for inclusion. These surveys are conducted approximately once every 2 years.

In addition, at regular intervals State licensure agencies, national voluntary associations, and other appropriate sources send to the Center their most recent directories or lists of new facilities. These lists are then clerically matched with the most current NMFI file, and facilities not already included are added.

For hospitals, data were gathered annually in a joint survey of the American Hospital Association and the Center. The contractual arrangement, which in effect merged the hospital portion of NMFI with the American Hospital Association's annual survey of hospitals, began in 1968. Through 1975 the American Hospital Association performed the data collection (from October through January each year) for its member hospitals, while the Center performed the data collection for the approximately 400 hospitals that were not registered with the American Hospital Association. The American Hospital Association processed both portions of the survey according to the same edit and processing specifications and delivered to the Center two edited tapes, one for American Hospital Association hospitals and one for nonmember hospitals. In 1976 and 1977 the American Hospital Association surveyed all hospitals as part of its annual survey. The contractual relationship with the American Hospital Association ceased in 1978. Since that time the Center has purchased hospital data tapes annually from the American Hospital Association, tapes which are solely for the Center's internal statistical use.

During the period 1976-80, under the health facilities component of the Cooperative Health Statistics System (CHSS), health facilities data for some States were collected and processed by the State CHSS contractor as a byproduct of the licensure renewal process. Edited data were then provided on computer tape by each CHSS State to the Center, and these data were merged into NMFI. Data for all facilities, except hospitals, in States not participating in CHSS continued to be collected by the Center every 2 years as part of NMFI. The next NMFI survey is planned for 1982.

*Content of data collection instruments.*—The fol-

lowing types of data are collected for the three categories of facilities found in the National Master Facility Inventory—hospitals, nursing homes, and other facilities:

*Hospitals*: Ownership; major type of service offered; whether various facilities and services are offered; number of beds, admissions, inpatient days of care, and discharges; patient census; number of bassinets, live births, and newborn days of care; outpatient utilization; number of surgical operations; revenue, expenses, and assets; and staffing.

*Nursing homes*: Ownership; major type of service; licensed and staffed beds; beds certified for Medicare and Medicaid; admission policy with regard to age, sex, and various conditions; patient census by age and sex; inpatient days of care; number of admissions, discharges, and deaths; staffing; who is in charge of nursing care; number of patients receiving nursing care; services routinely provided; basic monthly charge; and operating expenses.

*Other facilities*: Ownership; major type of service; licensed and staffed beds; beds certified as intermediate care beds; admission policy regarding age and sex; patient census by age and sex; inpatient days of care; number of admissions, discharges, and deaths; staffing; basic monthly charge; and operating expenses.

A copy of the data collection instruments is included in Series 14 reports based on NMFI data.

*Data processing and quality control.*—As already stated, NMFI is considered to be the most comprehensive file of inpatient health facilities available in the United States. In order to measure statistically the extent of its coverage, a "complement" survey was developed. It is an application of a general technique often called a "multiframe survey." In this application there are two frames—the National Master Facility Inventory and a geographic area sample list. From a probability area sample, all institutions found in the sample areas are identified, and the probability with which each comes into the sample is determined. Those inpatient health facilities found in the area sample survey (the complement survey) are matched against the NMFI list of facilities for that area. Any facility discovered within the scope of the sample survey but missed by NMFI constitutes undercoverage, and an appropriate weight is assigned to the omitted facility.

Data for nonresponding facilities, as well as data for questions not answered on the returned questionnaires, are obtained through various types of imputation procedures. Whenever possible, data from the previous NMFI survey are used to replace missing data. For those responding facilities that omitted data that are not available from a previous survey, an imputation method is used. In this method, all facilities of the same type are stratified by predetermined variables such as bed size groups, ownership categories, and major types of service. Missing data for a facility are then supplied by using the data from a facility of like characteristics. When a responding facility fails to answer any of the major questions (regarding, for example, bed size, ownership, type of facility, or number of patients), a followup questionnaire containing the questions omitted is sent to the respondents and they are asked again to answer.

*Release of data.*—Data from NMFI surveys are published in various Center reports. The biennial editions of *Health Resources Statistics* contain the most current available data (often preliminary estimates) from these surveys. Final data are published in Series 14 of *Vital and Health Statistics* and in two special reports (one for hospitals and one for nursing homes) showing data by county and standard metropolitan statistical area. Directories of facilities for the mentally retarded and of nursing homes have been published from NMFI survey data. Methodology reports appear in Series 1 and 2 of *Vital and Health Statistics*.

NMFI data are also released in the form of special tabulations prepared specifically for individual requests, computer tapes available through the Center's data tape release program, and special printouts or listings from the data tapes. Approximately 45-50 tape copies are purchased annually by such organizations as drug companies, State agencies, other Federal agencies, private contractors working on projects for other government agencies, universities, and national health care organizations.

Only the financial data from the hospital data tapes prior to 1977 are confidential; all other information is available for individual hospitals. However, beginning with the 1977 hospital data tape, the data release agreement between the American Hospital Association and the Center has limited the availability of hospital data. Portions of the nursing home and other health facility data are collected under a pledge of confidentiality and can be released only in statistical aggregates or with all identification items stripped.

### National health professions inventories and surveys

*Background.*—The National Center for Health Statistics first embarked on a health professions statistics program in 1964, when an ongoing program involving the compilation of secondary data on eight health occupations was transferred from the Division of Public Health Methods to the Center. As an adjunct to secondary source data, various surveys of facilities and health occupations were conducted in subsequent years. In 1974 the Center's health professions program was further enhanced by the transfer

of various surveys from the Bureau of Health Manpower (now the Bureau of Health Professions) to the Center.

*Purpose and scope.*—The primary purpose of the Center's health professions data collection is to produce reliable statistics on distribution; availability; and demographic, educational, and professional characteristics of persons employed in the health occupations. Under the health professions component of the Cooperative Health Statistics System for the 13 health occupations that are licensed in all of the 50 States and the District of Columbia, data were collected in the late 1970's on an inventory basis covering all licensees in participating States. The 13 health occupations covered were doctors of medicine, doctors of osteopathy, pharmacists, chiropractors, dentists, dental hygienists, registered nurses, licensed practical nurses, nursing home administrators, optometrists, physical therapists, podiatrists, and veterinarians. Data can be obtained on the nonlicensed health occupations largely through special surveys of facilities or of other employers. Special surveys are required because of the large number of employers of health workers, the limited use of licensure in the health field, the large and rapidly growing numbers and types of health workers, and the inadequate data base.

*Data collection procedures.*—Recent health professions inventories were conducted principally through the mechanism of the Cooperative Health Statistics System. Those States that participated in CHSS collected and processed the data according to specifications provided by the Center. They then provided the Center with data tapes. For those States not participating in CHSS, data for three professions (pharmacists, optometrists, and registered nurses) were collected by the national organizations representing these professions, under contract to the Center. The Center then merged data from these various sources to obtain national data for the three professions.

Data collection was performed in conjunction with each State's licensure renewal mechanism, either on an annual or on a biennial basis as dictated by the licensing procedures in each State. The licensure mechanism offers a number of distinct advantages for the development of a data system for the production of timely, accurate, and complete statistics on health professionals. Licensure data provide accurate counts and distributions by geographic areas of licensed health personnel and contain the necessary information for contacting all persons employed in a health profession in order to obtain additional detail.

Data were obtained by means of a self-administered questionnaire, which was either mailed out as an attachment to the application for license or enclosed as a separate form. Both mail and telephone followup procedures were used with nonrespondents. Approximately 10 minutes was required for the completion of a questionnaire. Item completion rates of at least 90 percent were generally achieved, and coverage approached 100 percent of the universe.

The national health professions surveys conducted by the Center have been of two types: censuses of facilities (such as the National Master Facility Inventory) to gather headcount data for detailed lists of health occupations employed in such facilities as hospitals, nursing homes, group practices, and clinics; and surveys of facilities to gather detailed information about specific types of health workers, such as nursing personnel in hospitals, health care administration personnel, and public health personnel.

The mail questionnaire represents the primary method of data collection for most of these surveys. Both mail and telephone followup procedures were used. Time required for completion of the questionnaire varied, depending on the content of the particular survey. Response to these surveys generally exceeds 80 percent.

*Content of questionnaire.*—For the health professions inventories conducted through CHSS, occupation-specific minimum data sets were established for some of the 13 occupations. Each minimum data set consisted of standard core items basic to all 13 health professions and additional data items unique to each specific occupation. The following items were generally included for all occupations:

*Administrative data*: Licensing State, occupation, type of licensure action, name, and mailing address of licensee.

*Education and training*: Location of school of graduation, year of graduation, and total years of education or highest degree.

*Demographic characteristics*: Date of birth, sex, marital status, race and ethnicity.

*Employment characteristics*: Occupational specialty, years active in the occupation, weeks worked in the past 12 months, current activity status, geographic location of place of practice 1 year ago, geographic location of current primary place of work, principal setting of primary activity, and number of hours usually spent per week by type of activity.

The Center's health professions surveys vary in content and purpose. At a minimum, the censuses of facilities obtain information on full- or part-time status and credentialing status. The sample surveys generally are designed to obtain data on personal characteristics, educational background, specialty, practice setting, and location.

*Data processing procedures.*—The data from the inventories and surveys are generally processed by the agency or organization responsible for the data

collection. Standardized coding, editing, and imputation procedures are developed by the Center and provided to the contractor for purposes of ensuring both the quality and comparability of the data. The Center unduplicates the edited data in order to account for persons licensed by more than one State. In addition, the Center performs considerable verification and cross-checking of the data against other lists and sources of health personnel data.

*Release of data.*—The unduplicated data are summarized and disseminated by the Center to all participating agencies in a manner that is comparable among State and local areas in terms of content, definitions, and format. The data appear in the following Center publications: *Vital and Health Statistics* Series 14; *Health Resources Statistics*; and various State reports. Special tabulations are available on request, and copies of many of the data tapes may be purchased. The data tapes are stripped of all identifiable information pertaining to individuals.

*Uses of data.*—Analysis of the data is intended to fulfill the following purposes: to assess current health personnel capabilities relative to the needs for health services at the local, State, and national levels; to project future demands on the supply of health professionals; to increase the efficiency of such data collection efforts; and to reduce the effort and costs involved in disseminating and using health professions information. The inventories can also be used as a sample frame for special surveys designed to obtain more detailed or specialized information.

# Appendixes

# Contents


I. Definition of certain terms used in this report ..... 31
- Terms relating to population and geography ..... 31
- Terms relating to health ..... 31
- Terms relating to sampling ..... 32
- Terms relating to data processing ..... 33

II. Listing of reports related to specific programs ..... 34
- Basic vital statistics ..... 34
- Vital statistics followback surveys ..... 35
- National Survey of Family Growth ..... 35
- National Health Interview Survey ..... 35
- National Health and Nutrition Examination Survey ..... 36
- National Hospital Discharge Survey ..... 37
- National Ambulatory Medical Care Survey ..... 37
- National Nursing Home Survey ..... 37
- National Master Facility Inventory ..... 37

**Appendix figure**

Regions and divisions of the United States as defined by the U.S. Bureau of the Census ..... 32

# Appendix I. Definition of certain terms used in this report

Brief definitions of some of the technical or specialized terms used in this report are provided. These are simplified definitions of what are often complex procedures, and the definitions are somewhat general in nature. For detailed definitions of the terms as they are used in the context of specific surveys, a reader should refer to the definitions contained in the appendix of the appropriate reports for each survey.

### Terms relating to population and geography

*Resident population.*—The resident population is the population living in the United States. This includes members of the Armed Forces stationed in the United States and their families as well as foreigners working or studying here; it excludes foreign military, naval, and diplomatic personnel and their families located here and residing in embassies or similar quarters as well as Americans living abroad. The resident population is often the denominator when calculating birth and death rates and incidence of disease.

*Civilian population.*—The civilian population is the resident population excluding members of the Armed Forces. Families of members of the Armed Forces are included, however.

*Civilian noninstitutionalized population.*—The civilian noninstitutionalized population is the civilian population not residing in institutions. Institutions include correctional institutions, detention homes, and training schools for juvenile delinquents; homes for the aged and dependent (for example, nursing homes and convalescent homes); homes for dependent and neglected children; homes and schools for the mentally or physically handicapped; homes for unwed mothers; and psychiatric, tuberculosis, and chronic disease hospitals and residential treatment centers.

*Household.*—All persons occupying a single housing unit are referred to as a household for sampling purposes.

*Family.*—A family is two or more persons living in the same household who are related by blood, marriage, or adoption. All persons living in a household who are related to each other are regarded as one family.

*Group quarters.*—Group quarters are quarters occupied by five or more persons unrelated to the head of the household; quarters with no designated head but with six or more unrelated persons are also group quarters. Also included are some quarters occupied by only one or two persons (for example, one to five persons occupying a surgical ward of a general hospital who have no usual residence elsewhere); also included are quarters occupied by persons in military barracks and college dormitories.

*Registration area.*—The United States has separate registration areas for birth, death, marriage, and divorce statistics, in which data are collected annually from States whose registration data are at least 90-percent complete.

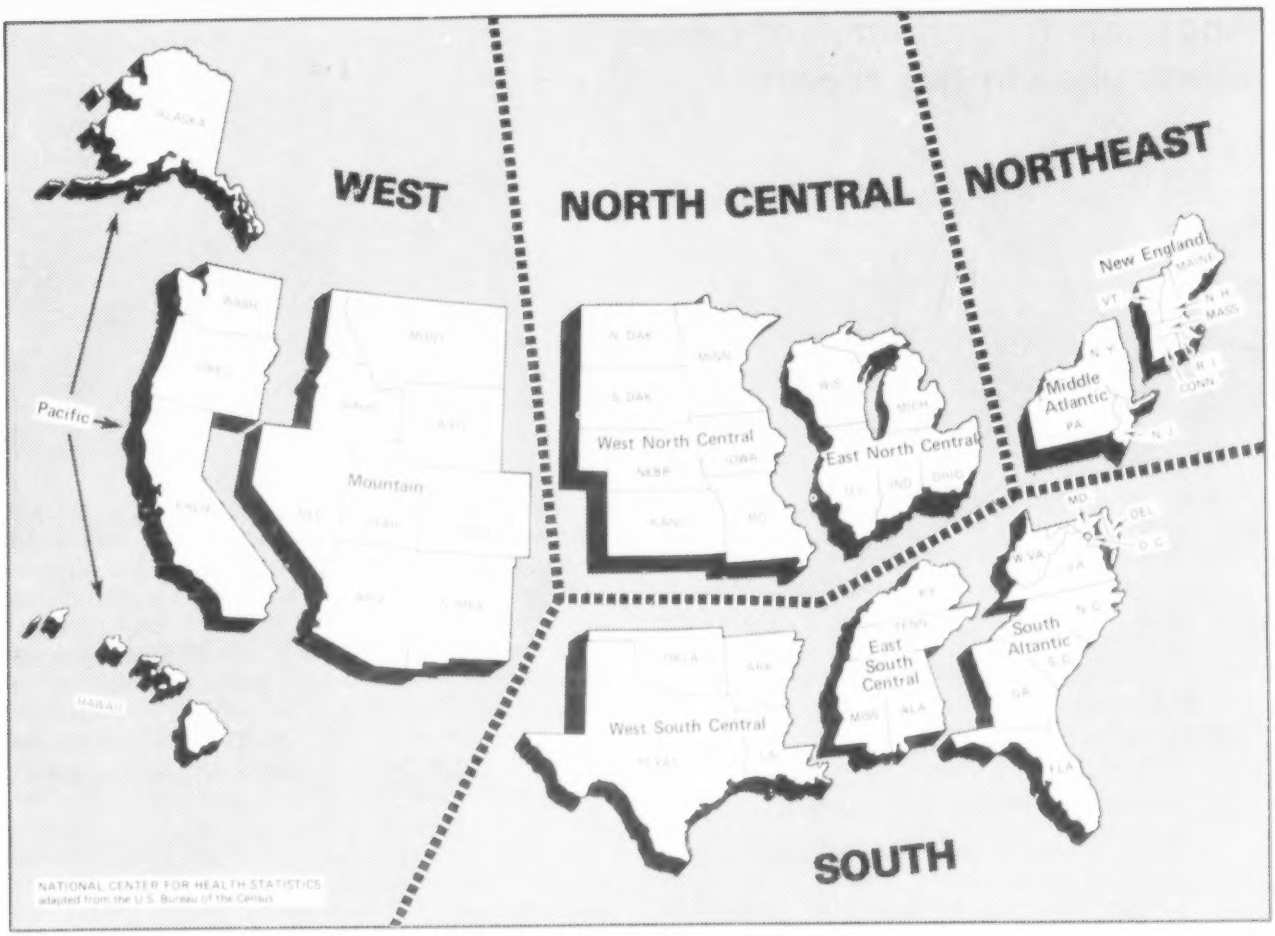
*Region and division.*—The 50 States and the District of Columbia are grouped for statistical purposes by the U.S. Bureau of the Census into nine divisions within four regions as shown on the map provided.

### Terms relating to health

*Condition.*—A health condition is a departure from a state of physical or mental well-being. Conditions are coded by the Center to the most recent revision of the International Classification of Diseases.

*Incidence.*—Incidence is the number of cases of disease having their onset during a prescribed period of time; it is often expressed as a rate (for example, the incidence of measles per 1,000 children 5-15 years of age during a year). Incidence is a measure of morbidity or other events that occur within a specified period of time.

*Prevalence.*—Prevalence is the number of cases of a disease, infected persons, or persons with some




Regions and divisions of the United States as defined by the U.S. Bureau of the Census

other attribute present during a particular interval of time. It is often expressed as a rate (for example, the prevalence of diabetes per 1,000 persons during a year).

*Medicare (Title XVIII).*—Medicare is a nationwide health insurance program providing health insurance protection to people 65 years of age and over, people eligible for social security disability payments for more than 2 years, and people with end-stage renal disease, regardless of income. The program was enacted July 30, 1965, as Title XVIII, *Health Insurance for the Aged*, of the Social Security Act, and became effective on July 1, 1966. It consists of two separate but coordinated programs: hospital insurance (Part A) and supplementary medical insurance (Part B).

*Medicaid (Title XIX).*—Medicaid is a federally aided but State operated and administered program. It provides medical benefits for certain low-income persons in need of medical care. The program, authorized in 1965 by Title XIX of the Social Security Act, categorically covers participants in the Aid to Families with Dependent Children program as well as some participants in the Supplemental Security Income program and other people deemed medically needy in a participating State. States also determine the benefits covered, rates of payment for providers, and methods of administering the program.

*Inpatient facility.*—An inpatient facility is, in general, one in which residents or patients are admitted for overnight stay for care. Included are hospitals, nursing homes, and various custodial and rehabilitation facilities.

*Outpatient facility.*—An outpatient facility is, in general, a place which treats patients without admitting them for an overnight stay. Included are physicians' offices, neighborhood health centers, and hospital outpatient departments.

### Terms relating to sampling

*Target population.*—The target population is the population about which information is wanted.

*Sampling frame.*—A sampling frame is a complete list of empirical cases that constitute the universe (or total population) of cases that correspond to the conceptual universe to which findings of a sample survey are generalized.

*Probability sample.*—The distinguishing characteristic of this type of sample is that every unit in the frame must have a known, but not necessarily equal, probability of being included in the sample.

*Stratified sample.*—All individual units in the sampling frame of a stratified sample are divided into groups or categories that are defined in such a way that each unit appears in one and only one stratum. Independent samples are selected within each group or stratum. The object is to arrange the population into strata that differ as widely as possible from each other but are as homogeneous as possible within each stratum.

*Area probability sample.*—An area probability sample is a procedure for demographic surveys of the population in which the initial frame consists not of individuals but of geographic areas. The sampling progresses through a series of stages in which first areas, then dwelling units, and finally the individuals themselves are selected.

*Primary sampling unit.*—In multistage samples, a primary sampling unit is the first level of sampling; in area probability sampling the United States is often divided into groups of counties, each group of which is a primary sampling unit.

*Sampling error.*—Sampling error is that part of the error of an estimate that is due to the fact that the estimate is obtained from a sample rather than from a 100-percent enumeration using the same procedures.

*Nonsampling error.*—Nonsampling error is the error in an estimate arising at any stage in a survey from such sources as varying interpretation of questions by enumerators, unwillingness or inability of respondents to give correct answers, nonresponse, improper coverage, and other sources exclusive of sampling error.

*Response rate.*—The response rate for a survey is the percent of an eligible sample for whom information is obtained. For an interview survey the numerator of the formula is the number of completed interviews. The denominator is the total sample size minus noneligible respondents, that is, minus those not meeting the criteria for a potential respondent as defined for a particular study.

### Terms relating to data processing

*Coding.*—Coding is a technical procedure for converting verbal information into numbers or other symbols that can be more easily counted and tabulated.

*Coding error.*—Coding error is the assignment of an incorrect code to a survey response.

*Editing.*—Editing is a preliminary step in data handling in which the responses are inspected, corrected, and sometimes precoded according to a fixed set of rules.

*Imputation.*—Imputation is the process of developing estimates for missing or inconsistent data in a survey. Data obtained from other units in the survey are usually used in developing the estimate.

*Quality control.*—Quality control is any observation or procedure used in any operation of a survey in order to prevent or reduce the effect of nonsampling errors. It includes many types of data editing and recoding.

# Appendix II. Listing of reports related to specific programs

This list of reports from the National Center for Health Statistics relating to each of the data collection programs is provided to facilitate identification of reports about each program; the list *excludes* analytic reports from the programs. Information about the availability of these publications can be obtained from the Center's Scientific and Technical Information Branch, Division of Data Services.

Copies of all Center reports are typically available at large municipal libraries, medical libraries, and libraries of schools of public health. Review of publications at these places would assist in determining whether a particular report is of interest.

**Basic vital statistics**

*Series 2*

No. 83 Comparability of Reporting Between the Birth Certificate and the National Natality Survey. 44 pp.

No. 81 A Methodological Study of Quality Control Procedures for Mortality Medical Coding. 40 pp.

No. 79 Vital Registration Systems in Five Developing Countries: Honduras, Mexico, Philippines, Thailand, and Jamaica. 159 pp.

No. 66 Comparability of Mortality Statistics for the Seventh and Eighth Revisions of the International Classification of Diseases, United States. 93 pp.

No. 46 Vital Signs Present at Birth. 19 pp.

No. 34 Comparison of Marital Status, Race, Nativity, and Country of Origin on the Death Certificate and Matching Census Record, United States, May-August 1960. 47 pp.

No. 30 Comparison of the Classification of Place of Residence on Death Certificates and Matching Census Records, United States, May-August 1960. 60 pp.

No. 29 Comparability of Age on the Death Certificate and Matching Census Record, United States, May-August 1960. 53 pp.

No. 20 Variance and Covariance of Life Table Functions Estimated From a Sample of Deaths. 8 pp.

No. 4 Comparison of Two Methods of Constructing Abridged Life Tables by Reference to a "Standard" Table. 11 pp.

*Series 3*

No. 19 National Estimates of Marriage Dissolution and Survivorship. 31 pp.

No. 18 Geographic Patterns in the Risk of Dying and Associated Factors, Ages 35-74 Years, United States, 1968-72. 120 pp.

No. 16 Cohort Mortality and Survivorship: United States Death Registration States, 1900-68. 36 pp.

No. 15 Trends in "Prematurity," United States, 1950-67. 51 pp.

No. 6 International Comparisons of Perinatal and Infant Mortality: The United States and Six West European Countries. 97 pp.

No. 4 Infant and Perinatal Mortality in the United States. 87 pp.

No. 1 The Change in Mortality Trend in the United States. 43 pp.

*Series 4*

No. 18 Statistics Needed for National Policies Related to Fertility. 154 pp.

No. 12 Needs for National Studies of Population Dynamics: A Report of the U.S. National Committee on Vital and Health Statistics. 31 pp.

No. 8 The 1968 Revision of the Standard Certificates. 47 pp.

No. 3 Report of the International Conference on the Perinatal and Infant Mortality Problem of the United States. 21 pp.

No. 2 National Vital Statistics Needs: A Report of the U.S. National Committee on Vital and Health Statistics. 51 pp.

No. 1 Fertility Measurement: A Report of the U.S. National Committee on Vital and Health Statistics. 26 pp.

*Vital Statistics of the United States Annual Volumes*

Volume I, Natality
Volume II, Mortality, Part A
Volume II, Mortality, Part B
Volume III, Marriage and Divorce

**Vital statistics followback surveys**

*Series 2*

No. 83 Comparability of Reporting Between the Birth Certificate and the National Natality Survey. 44 pp.

No. 56 Completeness and Quality of Response in the North Carolina Marriage Follow-Back Survey. 42 pp.

**National Survey of Family Growth**

*Series 2*

No. 87 National Survey of Family Growth, Cycle II: Sample Design, Estimation Procedures, and Variance Estimation. 36 pp.

No. 76 National Survey of Family Growth, Cycle I: Sample Design, Estimation Procedures, and Variance Estimation. 32 pp.

*Series 3*

No. 17 Patterns of Aggregate and Individual Changes in Contraceptive Practice, United States, 1965-1975. 23 pp.

*Series 4*

No. 18 Statistics Needed for National Policies Related to Fertility. 154 pp.

No. 12 Needs for National Studies of Population Dynamics: A Report of the U.S. National Committee on Vital and Health Statistics. 31 pp.

**National Health Interview Survey**

*Series 1*

No. 11 Health Interview Survey Procedure, 1957-74. 153 pp.

No. 2 Health Survey Procedure: Concepts, Questionnaire Development, and Definitions in the Health Interview Survey. 66 pp.

*Series 2*

No. 82 Small Area Estimation: An Empirical Comparison of Conventional and Synthetic Estimators for States. 19 pp.

No. 75 Synthetic Estimation of State Health Characteristics Based on the Health Interview Survey. 22 pp.

No. 69 A Summary of Studies of Interviewing Methodology. 78 pp.

No. 57 Net Differences in Interview Data on Chronic Conditions and Information Derived From Medical Records. 58 pp.

No. 54 Quality Control and Measurement of Nonsampling Error in the Health Interview Survey. 53 pp.

No. 52 Reliability of Estimates with Alternate Cluster Sizes in the Health Interview Survey. 17 pp.

No. 50 Optimum Recall Period for Reporting Persons Injured in Motor Vehicle Accidents. 31 pp.

No. 49 Reporting Health Events in Household Interviews: Effects of an Extensive Questionnaire and a Diary Procedure. 80 pp.

No. 48 Interviewing Methods in the Health Interview Survey. 86 pp.

No. 45 Reporting Health Events in Household Interviews: Effects of Reinforcement, Question Length, and Reinterviews. 70 pp.

No. 41 Effect of Some Experimental Interviewing Techniques on Reporting in the Health Interview Survey. 62 pp.

No. 38 Estimation and Sampling Variance in the Health Interview Survey. 28 pp.

No. 37 Development and Evaluation of an Expanded Hearing Loss Scale Questionnaire. 42 pp.

No. 26 The Influence of Interviewer and Respondent Psychological and Behavioral Variables on the Reporting in Household Interviews. 65 pp.

No. 23 Interview Data on Chronic Conditions Compared With Information Derived From Medical Records. 84 pp.

No. 18 Interview Response on Health Insurance Compared With Insurance Records, United States, 1960. 43 pp.

No. 16 Identifying Problem Drinkers in a Household Health Survey. 46 pp.

No. 12 Methodological Aspects of a Hearing Ability Interview Survey. 19 pp.

No. 10 Hospital Utilization in the Last Year of Life. 30 pp.

No. 8 Comparison of Hospitalization Reporting in Three Survey Procedures. 48 pp.

No. 7 Health Interview Responses Compared With Medical Records. 74 pp.

No. 6 Reporting of Hospitalization in the Health Interview Survey. 71 pp.

No. 2 Measurement of Personal Health Expenditures. 59 pp.

*Series 3*

No. 7 Selected Family Characteristics and Health Measures Reported in the Health Interview Survey. 26 pp.

**National Health and Nutrition Examination Survey**

*Series 1*

No. 15 Plan and Operation of the Second National Health and Nutrition Examination Survey, 1976-80. 144 pp.

No. 14 Plan and Operation of the HANES I Augmentation Survey of Adults 25-74 Years, United States, 1974-75. 110 pp.

No. 10 Plan and Operation of the Health and Nutrition Examination Survey, United States, 1971-73. Part A, 46 pp. Part B, 77 pp.

No. 8 Plan and Operation of a Health Examination Survey of U.S. Youths 12-17 Years of Age. 80 pp.

No. 5 Plan, Operation, and Response Results of a Program of Children's Examinations. 56 pp.

No. 4 Plan and Initial Program of the Health Examination Survey. 43 pp.

*Series 2*

No. 86 Computer Assisted Spirometry Data Analysis for the National Health and Nutrition Examination Survey. 48 pp.

No. 74 The Construction and Utility of Three Indexes of Intellectual Achievement—An Intellectual-Development (ID) Index, a Socio-Intellectual-Status (SIS) Index, a Differential-Intellectual-Development(DID) Index—U.S. Children and Youths, 6-17 Years. 26 pp.

No. 73 A Concurrent Validational Study of the NCHS Well-Being Schedule. 53 pp.

No. 72 Methodologic Programs in Children's Spirometry. 43 pp.

No. 71 Development of the Speech Reception Test. 16 pp.

No. 67 A Study of the Effect of Remuneration Upon Response in the Health and Nutrition Examination Survey, United States. 23 pp.

No. 62 Language and Adjustment Scales for the Thematic Apperception Tests for Youths 12-17 Years. 84 pp.

No. 60 The Rationale, Development, and Standardization of a Basic Word Vocabulary Test. 71 pp.

No. 59 Vision Test Validation Study for the Health Examination Survey Among Youths 12-17 Years. 36 pp.

No. 58 Language and Adjustment Scales for the Thematic Apperception Test for Children 6-11 Years. 70 pp.

No. 47 Subtest Estimates of the WISC Full-Scale IQ's for Children. 42 pp.

No. 44 Quality Control in a National Health Examination Survey. 22 pp.

No. 43 Sample Design and Estimation Procedures for a National Health Examination Survey of Children. 40 pp.

No. 40 Loudness Balance Study of Selected Audiometer Earphones. 37 pp.

No. 36 Factors Related to Response in a Health Examination Survey, United States, 1960-62. 48 pp.

No. 35 Comparison of Timed and Untimed Presentation of the Goodenough-Harris Test of Intellectual Maturity. 16 pp.

No. 27 Development of the Brief Test of Literacy. 29 pp.

No. 25 Orthodontic Treatment Priority Index. 49 pp.

No. 24 A Study of the Achievement Test Used in the Health Examination Surveys of Persons Aged 6-17 Years. 60 pp.

No. 22 Three Views of Hypertension and Heart Disease. 43 pp.

No. 21 Calibration of Two Bicycle Ergometers Used by the Health Examination Survey. 10 pp.

No. 15 Evaluation of Psychological Measures Used in the Health Examination Survey of Children Ages 6-11. 67 pp.

No. 9 Cooperation in Health Examination Surveys. 38 pp.

No. 3 The One-Hour Oral Glucose Tolerance Test. 34 pp.

No. 1 Comparison of Two Vision-Testing Devices. 33 pp.

*Series 3*

No. 14 Height and Weight of Children in the United States, India, and the United Arab Republic. 51 pp.

**National Hospital Discharge Survey**

*Series 2*

No. 88 The Status of Hospital Discharge Data in Denmark, Scotland, West Germany, and the United States. 58 pp.

No. 80 The Status of Hospital Discharge Data in Six Countries. 77 pp.

No. 68 Quality Control in the Hospital Discharge Survey. 41 pp.

No. 55 The Prediction Approach to Finite Population Theory: Application to the Hospital Discharge Survey. 31 pp.

No. 39 Development of the Design of the NCHS Hospital Discharge Survey. 29 pp.

No. 28 Pilot Study on Patient Charge Statistics. 59 pp.

No. 19 Participation of Hospitals in the Pilot Study of the Hospital Discharge Survey. 55 pp.

No. 13 Computer Simulation of Hospital Discharges. 34 pp.

*Series 4*

No. 14 Uniform Hospital Abstract, Minimum Basic Data Set: A Report of the U.S. National Committee on Vital and Health Statistics. 17 pp.

No. 11 Use of Hospital Data for Epidemiologic and Medical-Care Research: A Report of the U.S. National Committee on Vital and Health Statistics. 9 pp.

**National Ambulatory Medical Care Survey**

*Series 2*

No. 78 A Reason for Visit Classification for Ambulatory Care. 63 pp.

No. 63 The National Ambulatory Medical Care Survey: Symptom Classification. 35 pp.

No. 61 National Ambulatory Medical Care Survey: Background and Methodology, United States, 1967-72. 76 pp.

*Series 4*

No. 16 Ambulatory Medical Care Records, Uniform Minimum Basic Data Set: A Report of the U.S. National Committee on Vital and Health Statistics. 16 pp.

**National Nursing Home Survey**

*Series 1*

No. 7 Design and Methodology for a National Survey of Nursing Homes. 37 pp.

**National Master Facility Inventory**

*Series 1*

No. 12 Development of the National Inventory of Family Planning Services, United States. 43 pp.

No. 9 Design and Methodology of the 1967 Master Facility Inventory Survey. 30 pp.

No. 6 The Agency Reporting System for Maintaining the National Inventory of Hospitals and Institutions. 19 pp.

No. 3 Development and Maintenance of a National Inventory of Hospitals and Institutions. 25 pp.

blank page

# Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures.**—Reports describing the general programs of the National Center for Health Statistics and its offices and divisions and the data collection methods used. They also include definitions and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research.**—Studies of new statistical methodology including experimental tests of new survey methods, studies of vital statistics collection methods, new analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory.

SERIES 3. **Analytical and Epidemiological Studies.**—Reports presenting analytical or interpretive studies based on vital and health statistics, carrying the analysis further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports.**—Final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

SERIES 10. **Data from the National Health Interview Survey.**—Statistics on illness, accidental injuries, disability, use of hospital, medical, dental, and other services, and other health-related topics, all based on data collected in the continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey and the National Health and Nutrition Examination Survey.**—Data from direct examination, testing, and measurement of national samples of the civilian noninstitutionalized population provide the basis for (1) estimates of the medically defined prevalence of specific diseases in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics and (2) analysis of relationships among the various measurements without reference to an explicit finite universe of persons.

SERIES 12. **Data From the Institutionalized Population Surveys.**—Discontinued in 1975. Reports from these surveys are included in Series 13.

SERIES 13. **Data on Health Resources Utilization.**—Statistics on the utilization of health manpower and facilities providing long-term care, ambulatory care, hospital care, and family planning services.

SERIES 14. **Data on Health Resources: Manpower and Facilities.**—Statistics on the numbers, geographic distribution, and characteristics of health resources including physicians, dentists, nurses, other health occupations, hospitals, nursing homes, and outpatient facilities.

SERIES 15. **Data From Special Surveys.**—Statistics on health and health-related topics collected in special surveys that are not a part of the continuing data systems of the National Center for Health Statistics.

SERIES 20. **Data on Mortality.**—Various statistics on mortality other than as included in regular annual or monthly reports. Special analyses by cause of death, age, and other demographic variables; geographic and time series analyses; and statistics on characteristics of deaths not available from the vital records based on sample surveys of those records.

SERIES 21. **Data on Natality, Marriage, and Divorce.**—Various statistics on natality, marriage, and divorce other than as included in regular annual or monthly reports. Special analyses by demographic variables; geographic and time series analyses; studies of fertility; and statistics on characteristics of births not available from the vital records based on sample surveys of those records.

SERIES 22. **Data From the National Mortality and Natality Surveys.**—Discontinued in 1975. Reports from these sample surveys based on vital records are included in Series 20 and 21, respectively.

SERIES 23. **Data From the National Survey of Family Growth.**—Statistics on fertility, family formation and dissolution, family planning, and related maternal and infant health topics derived from a periodic survey of a nationwide probability sample of ever-married women 15-44 years of age.

For a list of titles of reports published in these series, write to:

Scientific and Technical Information Branch
National Center for Health Statistics
Public Health Service
Hyattsville, Md. 20782